ABSTRACT

Title of dissertation: APPLICATIONS OF ARTIFICIAL NEURAL NETWORKS IN LEARNING QUANTUM SYSTEMS

Ruizhi Pan
Doctor of Philosophy, 2023

Dissertation directed by: Professor Charles W. Clark
Department of Physics

Quantum machine learning is an emerging field that combines techniques in the disciplines of machine learning (ML) and quantum physics. Research in this field takes three broad forms: applications of classical ML techniques to quantum physical systems, quantum computing and algorithms for classical ML problems, and new ideas inspired by the intersection of the two disciplines. We mainly focus on the power of artificial neural networks (NNs) in quantum-state representation and phase classification in this work.

In the first part of the dissertation, we study NN quantum states which are used as wave-function ansätze in the context of quantum many-body physics. While these states have achieved success in simulating low-lying eigenstates and short-time unitary dynamics of quantum systems and efficiently representing particular states such as those with a stabilizer nature, more rigorous quantitative analysis about their expressibility and complexity is warranted. Here, our analysis of the restricted Boltzmann machine (RBM) state

representation of one-dimensional (1D) quantum spin systems provides new insight into their computational complexity. We define a class of long-range-fast-decay (LRFD) RBM states with quantifiable upper bounds on truncation errors and provide numerical evidence for a large class of 1D quantum systems that may be approximated by LRFD RBMs of at most polynomial complexities. These results lead us to conjecture that the ground states of a wide range of quantum systems may be exactly represented by LRFD RBMs or a variant of them, even in cases where other state representations become less efficient. At last, we provide the relations between multiple typical state manifolds. Our work proposes a paradigm for doing complexity analysis for generic long-range RBMs which naturally yields a further classification of this manifold. This paradigm and our characterization of their nonlocal structures may pave the way for understanding the natural measure of complexity for quantum many-body states described by RBMs and are generalizable for higher-dimensional systems and deep neural-network quantum states.

In the second part, we use RBMs to investigate, in dimensions $D = 1$ and 2, the many-body excitations of long-range power-law interacting quantum spin models. We develop an energy-shift method to calculate the excited states of such spin models and obtain a high-precision momentum-resolved low-energy spectrum. This enables us to identify the critical exponent where the maximal quasiparticle group velocity transits from finite to divergent in the thermodynamic limit numerically. In $D = 1$, the results agree with an analysis using the field theory and semiclassical spin-wave theory. Furthermore, we generalize the RBM method for learning excited states in nonzero-momentum sectors from 1D to 2D systems. At last, we analyze and provide all possible values ($\frac{3}{2}$, 2 and 3) of the critical exponent for 1D generic quadratic bosonic and fermionic Hamiltonians with long-range hoppings and

pairings which serves for understanding the speed of information propagation in quantum systems.

In the third part, we study deep NNs as phase classifiers. We analyze the phase diagram of a 2D topologically nontrivial fermionic model Hamiltonian with pairing terms at first and then demonstrate that deep NNs can learn the band-gap closing conditions only based on wave-function samples of several typical energy eigenstates, thus being able to identify the phase transition point without knowledge of Hamiltonians.

APPLICATIONS OF ARTIFICIAL NEURAL NETWORKS IN LEARNING QUANTUM SYSTEMS

by

Ruizhi Pan

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2023

Advisory Committee:
Professor Mohammad Hafezi, Chair
Professor Charles W. Clark, Co-Chair/Advisor
Professor Victor Yakovenko
Professor Alexey Gorshkov
Professor Christopher Jarzynski, Dean's representative

Acknowledgments

I would first like to express my heartfelt gratitude to my advisor Charles Clark, who has supported me throughout my Ph.D. study with his great patience and encouragement. I really appreciate the freedom that Charles gave to me in choosing research topics and exploring them at my own pace. I am really grateful for his instructions on how to delve into a specific research area and the so-called paper-torture session through which we could have a detailed discussion on manuscript writing. In fact, it is not torture for me and I really enjoy and cherish such a process. Working under the instruction of a knowledgeable, prestigious and amiable professor is my long-cherished wish and I think it is fulfilled at the University of Maryland, College Park.

I would like to thank many professors at the Joint Quantum Institute, the Joint Center for Quantum Information and Computer Science and other research institutes who offered me tremendous help in my Ph.D. research, including Alexey Gorshkov, Victor Yakovenko, Mohammad Hafezi, Andrew Childs, Ian Spielman, Sankar Das Sarma, Jay Deep Sau, Ana Maria Rey, Christopher Jarzynski, Jacob Taylor, Chris Greene, Ricardo Nochetto, Xiaoji Zhou, and Xuzong Chen. I really appreciate their willingness to share their knowledge and enthusiasm in science.

I also want to thank my friends in the department of physics, mathematics and computer science: Wenbo Li, Fangli Liu, Yi-Hsieh Wang, Yiming Cai, Peizhi Du, Bin Cao,

Yuchen Yue, Tongyang Li, Linfeng Zhang, Dong-ling Deng, Yidan Wang, Zhexuan Gong, Renxiong Wang, Xunnong Xu, Haitan Xu, Chunxiao Liu, Alejandra Maldonado-Trapp, Ben Eller, Tengfei Su, Peng Zhang and so many other colleagues. Their support and encouragement give me warmth and courage.

At last, I would like to thank my parents. Their love and support are my permanent source of strength to explore the world.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

BEC	Bose-Einstein Condensate
CQED	Cavity Quantum Electrodynamics
EM	Electromagnetic
LRFD	Long-range-fast-decay
MD	Multi-fold Degeneracy
MF	Majorana Fermion
ML	Machine Learning
MPS	Matrix Product State
NN	Neural Network
NNN	Next Nearest Neighbor
RBM	Restricted Boltzmann Machine
SC	Superconducting
SOC	Spin-orbit Coupling
TFIM	Transverse Field Ising Model
TRS	Time-Reversal Symmetry

Chapter 1: Introduction

1.1 Background

Quantum machine learning is an emerging field that combines techniques in the disciplines of machine learning (ML) and quantum physics [1, 2, 3, 4, 5, 6, 7]. Research in this field takes three broad forms [2]: applications of classical ML techniques to quantum physical systems [4, 5, 6, 7, 8, 9, 10], quantum computing and algorithms for classical ML problems [11, 12, 13, 14], and new ideas inspired by the intersection of the two disciplines [15, 16]. In the field of learning quantum systems, there has been tremendous progress in applying ML techniques to identifying quantum phases and transitions [3, 17, 18, 19, 20, 21], molecular modeling [22, 23], quantum state tomography [24, 25], and accelerating Monte Carlo simulations [26, 27]. While ML encompasses a wide range of modeling tools and computational algorithms to suit different needs in theoretical modeling and information processing, artificial neural networks (NNs), which are computing systems with specific architectures analogous to and actually inspired by the biological neural networks in animal brains, often play an important role due to their tremendous power in function approximation, classification and data processing.

One rapidly advancing field in recent years is the investigation of neural-network quantum states in the context of quantum many-body physics [5, 6, 7, 8, 9, 28, 29, 30, 31].

The core idea in this field is to postulate an ansatz for the wave function in terms of a neural network (NN) [6], which targets a low-dimensional manifold in the exponentially large Hilbert space for state approximation [32], and apply ML algorithms to find a specific solution. One of the most commonly used neural networks is the restricted Boltzmann machine (RBM) [6, 7, 8, 28, 29], which is a bipartite stochastic construct that combines the concepts of thermodynamic partition functions with those of classical artificial neural networks. The RBM usually works as the building block for understanding and training deeper networks because of its relatively simple structure for inference and its power in parametric modeling as a universal approximator for discrete distribution [33]. It has achieved success in representing a wide range of quantum states such as low-lying eigenstates of quantum many-body-localized systems [6, 34], code words of a stabilizer code [30, 35, 36] and chiral topological states [8, 37, 38].

One of the central challenges in state representation theory for quantum many-body states is to find efficient representations of states based on which the physical quantities of the global quantum systems can be extracted with information loss as less as possible [32]. It has been demonstrated that the RBMs as a representative of NNs have nonnegligible advantages in state representation for particular sets of quantum states including the states with entanglement entropy violating an area law, states of high-dimensional quantum systems and states related to chiral topological order [7]. But it remains to be determined what the natural measure of the complexity for quantum many-body states described by RBMs is and how to quantitatively study the expressibility of this state manifold.

To fully understand and exploit the power of NN state representation, we apply it to the investigation of long-range interacting quantum systems—those with interactions

decaying as a power law $1/r^{\alpha}$ in distance r. The low-lying eigenstates of such systems can possess quantum correlations with a long-range decaying and entanglement entropy violating the area law, thus are often associated with a higher parameterization complexity. In the past years, numerous atomic, molecular, and optical systems exhibiting long-range interactions are emerging as versatile platforms for quantum computation and quantum simulation [39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52]. So the investigation of the application of NNs to these systems can possibly benefit understanding quantum states with more complexity and provide potential benchmarks for large-scale near-term quantum simulators.

Besides the use in parameterizing wave functions for quantum systems, NNs with more complicated architectures, such as convolutional NNs, can also benefit the field of statistical physics in other ways. One of which is to use deep NNs to identify phases and phase transitions based on their ability to detect multiple types of order parameters, including some nontrivial states without conventional order. It is shown that NNs can identify phase transitions in correlated many-body systems without the knowledge of locality conditions of Hamiltonians [3]. We are curious about how NNs behave in learning the phase-transition points of quantum models with pairing terms.

1.2 Outline of dissertation

In this dissertation, we focus on quantitatively analyzing the expressibility and complexity of NN quantum states in state representation in the context of quantum many-body physics, investigating their application to learning many-body excitations of long-range

power-law interacting quantum spin models and using deep NNs for phase classification.

In Chapter 2, we study the efficiency of RBMs in state representation and provide a paradigm for doing complexity analysis for generic long-range RBMs. We propose a new concept—long-range-fast-decay (LRFD) RBM states with a quantified nonlocal structure. We derive an upper bound on truncation errors associated with two measures of state differences. Then we identify distinct asymptotic scaling laws of spatial complexities for RBMs determined by the nonlocal interaction pattern between physical and virtual particles in their forms. Based on this analysis, we provide numerical results supporting that ground states of a wide range of 1D quantum systems may be approximated by LRFD RBMs with at most polynomial complexities. These results offer evidence for the potentially high efficiency of RBMs in the scope where other state parameterizations become less efficient. At last, we provide the relations between multiple typical state manifolds.

In Chapter 3, we use RBMs to investigate the many-body excitations of long-range power-law interacting quantum spin models in dimensions $D = 1$ and 2. We develop an energy-shift method to calculate the excited states of such spin models and obtain a high-precision momentum-resolved low-energy spectrum. We numerically identify the critical exponent where the maximal quasiparticle group velocity transits from finite to divergent in the thermodynamic limit. In $D = 1$, we show that the results agree with an analysis using the field theory and semiclassical spin-wave theory. Furthermore, we generalize the RBM method learning excited states in nonzero-momentum sectors from 1D to 2D systems. At last, we analyze and provide all possible values ($\frac{3}{2}$, 2 and 3) of the critical exponent for generic 1D quadratic bosonic and fermionic Hamiltonians with long-range hoppings and pairings.

In Chapter 4, we study deep NNs as phase classifiers. We design a 2D topologically nontrivial fermionic model Hamiltonian with pairing terms and study its phase diagram at first and then demonstrate that deep NNs can learn the band-gap closing conditions only based on wave-function samples of several typical energy eigenstates (zero-energy edge modes), thus being able to identify the phase-transition point without knowledge of Hamiltonians.

Chapter 2: Efficiency of neural-network state representations of one-dimensional quantum spin systems

2.1 Introduction

In recent years, considerable activity has been devoted to the investigation of neural network quantum states in the context of quantum many-body physics [5, 6, 7, 8, 9, 28, 29, 30, 31]. The core idea is to postulate an ansatz for the wave function in terms of a neural network (NN) [6], which targets a low-dimensional manifold in the exponentially large Hilbert space for state approximation [32], and apply ML algorithms to find a specific solution. The restricted Boltzmann machine (RBM) [6, 7, 8, 28, 29] is a bipartite stochastic construct that combines the concepts of thermodynamic partition functions with those of classical artificial neural networks. The RBM usually works as the building block for understanding and training deeper networks because of its relatively simple structure for inference and its power in parametric modeling as a universal approximator for discrete distribution [33]. It has achieved success in representing a wide range of quantum states such as low-lying eigenstates of quantum many-body-localized systems [6, 34], code words of a stabilizer code [30, 35, 36] and chiral topological states [8, 37, 38].

While RBMs have demonstrated their power in numerical simulation, we have strong

motivations to theoretically and quantitatively investigate the expressibility and complexity of generic long-range RBMs. Here, long-range RBMs refer to RBMs with full connectivity between the visible (spin) and hidden layers, whose structural information is usually characterized by dense networks, in contrast to the so-called short-range or sparse RBMs [7, 30, 31].

The first motivation is that there is a high probability that the exact RBM representation of generic target quantum states naturally has a long-range form and needs at least exponentially many parameters. This argument is based on the empirical fact that most RBMs returned by ML algorithms after a full training for learning specific quantum states have dense network structures and the magnitudes of parameters characterizing connectivity generally decay but do not vanish as the site difference between visible and hidden nodes increases. The more and more hidden nodes added to the network, characterizing increasing computational resources, are able to capture correlations of higher and higher orders between spins [6].

Another piece of evidence is that, in the field of learning ground states of quantum systems, only for very few Hamiltonian families, such as those with a stabilizer nature, have we found a succinct RBM form to exactly represent the ground states [30, 31]. The difficulty of finding exact RBM solutions originates from the difficulty of reducing the number of nonlinear equations for parameters from exponential to polynomial or even linear in system sizes and solving them. Due to the high nonlinearity of RBM forms, it is possible that at least exponentially many hidden nodes are needed in an exact ground-state solution.

The second motivation is that the RBM solutions provided by relevant ML algorithms

are approximators of the exact target states and these approximations often feature the long-range form and the fast parameter decay as described above, even though the exact RBM representations of the target states are unknown or less efficient, or do not have such features. One example is the demonstration of high accuracy and less complexity for the approximated RBM representations of Jastrow wave functions obtained by numerical computations in the branch of learning chiral topological states [8]. Their exact RBM constructions can be derived in principle but have a polynomial scaling with a comparatively larger exponent, thus being less economical.

Moreover, we also find that the RBM construction for a target state, such as the ground state of a given Hamiltonian, is not unique in some cases even with a fixed global phase which implies eliminating the degree of freedom associated with a global gauge transformation. This argument can be justified using an example where the ground state of a spin-$\frac{1}{2}$ system is the state with all spins up. Therefore, if the algorithm for obtaining an RBM approximator has a stochastic nature [6], to which RBM form the solution converges may depend on the initial and hyperparameter settings.

As shown in many numerical works [6, 8], the RBM approximators in a wide range of cases hold the long-range and fast-decay features and have a form similar to a finite truncation of RBM forms with infinitely many parameters removing those with small enough magnitudes. Magnitude-based prunings can also be conducted for RBMs with a finite number of parameters and their effects may be phase and system dependent [53, 54]. Thus, it is worthwhile to generalize the form of the RBM wave-function ansatz by including infinitely many hidden nodes, quantitatively analyze the effects of finite truncations and justify the faithfulness of using these truncated long-range RBM approximators in math-

ematics. The core idea in viewing a conventional long-range RBM as a truncation of an "infinitely large" RBM is analogous to analyzing a finite truncation of the Taylor series in approximating an analytic function at a specific point while the limit of the infinite sequence of the Taylor polynomials converges to the exact value.

The third motivation for studying long-range RBMs stems from the central goal of exploring effective compressed state representations. It includes understanding the natural measure of complexity for quantum states described by a specific representation [7] and understanding how the global information and physical properties of the states are encoded in that description, just as how researchers interpret the tensor network states [32]. Additionally, there has been some work studying the relationship between RBMs and other concepts about state representations, such as string-bond states [8], correlator product states [55] and tensor network states [38, 56]. Especially, the transformation from RBMs to matrix product states (MPSs) [56] provides one way in principle to analyze the spatial complexity of RBMs. But such types of transformations may lead to redundancy when the RBMs are not quite short-range or sparse as they only use structural information of the network. It is perhaps also inconvenient to analyze the effects of truncations applied to RBMs through an extra intermediate transformation to other representations.

Furthermore, the RBM is an architecture that can naturally describe quantum states in a nonlocal manner [7], which is not the same as tensor network states that aim to encode the global wave-function information into a local tensor operator [32]. Therefore, it is strongly desirable to find a way to analyze the spatial complexity and extract information about the physical properties of long-range RBMs themselves.

In this work, we analyze the efficiency of long-range RBM state representation for 1D

quantum spin systems. Our procedure is as follows:

1. In Sec. 2.2.1, we generalize the RBM wave-function ansatz to an infinitely-many-hidden-node regime and define a subset of generic RBM states—the long-range-fast-decay (LRFD) RBM states, whose parameter conditions constrain the nonlocal interactions between spins (visible nodes) and virtual particles (hidden nodes).

2. In Sec. 2.2.2, we derive an upper bound on truncation errors associated with two measures of state differences for the sequence of truncated LRFD RBM states. One measure is the l_2-norm of the state-vector difference and the other is a Hermitian-operator-based expectation-value difference.

3. In Sec. 2.2.3, we identify the dependence of the spatial complexity for LRFD RBMs in state approximation on the decaying rates specified in the nonlocal interaction pattern.

4. In Sec. 2.3, we provide numerical evidence supporting a conjecture that the ground states of a wide range of 1D quantum spin systems, including some critical systems with logarithmic entanglement entropy, can be approximated by LRFD RBMs with the scaling of the spatial complexity being at most polynomial in both the system size and the inverse of approximation errors.

5. In Sec. 2.3, we also provide the relations between multiple typical state manifolds through which the importance of the concept of LRFD RBMs in efficiency analysis for state representation theory is manifested.

Our results offer evidence for the utility of RBMs in cases where other state parame-

terizations, such as matrix product states (MPSs), become less efficient. Our work actually proposes a paradigm of doing complexity analysis for general long-range RBMs, rather than limited to short-range or sparse RBMs, and naturally yields a further classification of this manifold based on the complexity scaling.

We find that the nonlocal structure of LRFD RBMs can be characterized by two conditions. These conditions are each determined by bounds associated with two degrees of freedom, defined within a framework of *levels* that is depicted in Fig. 2.1(a). One of the two degrees of freedom is a single-level decaying factor resembling localized orbitals and encoding information about correlations between spins (Sec. 2.2.4). The second is a level-decay factor, which has a significant influence on the complexity of the RBMs (Sec. 2.2.3).

This paradigm and our characterization of the nonlocal structures may promote the understanding of the natural measure of complexities for quantum many-body states described by RBMs and may be generalizable to higher-dimensional systems and to deep neural-network quantum states.

2.2 The restricted Boltzmann machine as a wave-function ansatz

We use the RBM as a wave-function ansatz for 1D quantum many-body spin-$\frac{1}{2}$ systems [5, 6, 7]. The RBM usually works as the building block for understanding and training deeper networks because of its relatively simple structure for inference and its power in parametric modeling as a universal approximator for discrete distribution [33]. As basic constructs of deep NNs, the RBMs have two *layers*. The first layer (a *visible* layer) represents a spin configuration $\vec{\sigma}$ in the usual way. Here, the vector $\vec{\sigma} = (\sigma_1, \ldots, \sigma_L)$

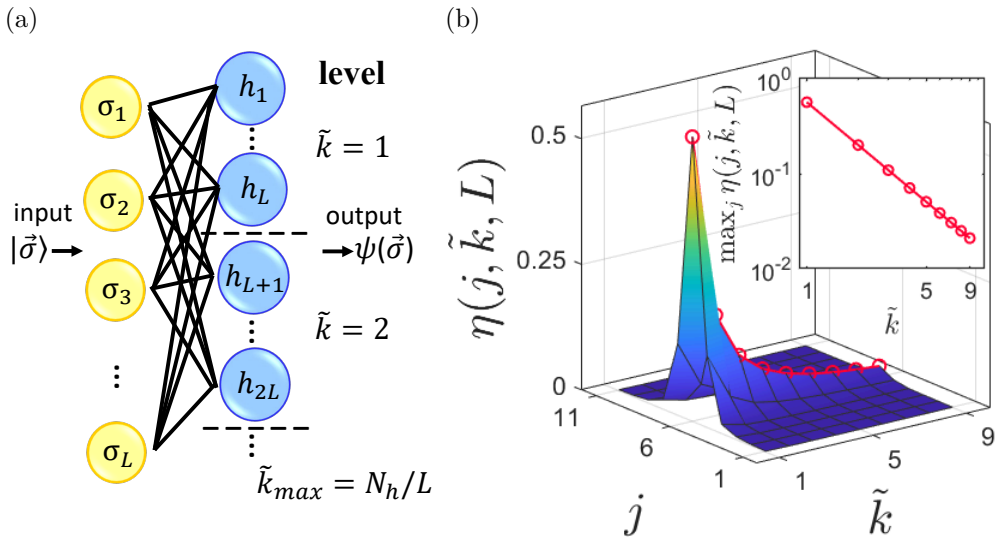


Figure 2.1: (a) Network structure of RBMs as a wave-function ansatz for 1D quantum spin systems. Long-range RBMs usually implies full connectivity between the visible and hidden layer. The hidden layer is divided into N_h/L levels, each containing L hidden nodes. (b) Importance measure $\eta(j, \tilde{k}, L)$ for a LRFD RBM with translational symmetry. The RBM is constructed as Eqs. (2.24)–(2.26) show, where $\lambda(\tilde{k}) = \tilde{k}^{-\alpha_P}$, $\mu(r) = \frac{1}{2}\delta_Q r^{-\alpha_Q}$ for $r \neq 0$, $\mu(0) = \delta_Q = 0.5$, $\alpha_P = 0.75$, $\alpha_Q = 1.5$, $c_w = 1 + i$, $c_b = 0$, $a_0 = 0$ and $L = 11$. The inset plots the decay of the maximum of $\eta(j, \tilde{k}, L)$ among all j at each level with increasing $\tilde{k}$ on a log-log scale. The linearity of the curve reveals a power-law decaying of the "ridge" (red circles) of the 3D structure.

represents a system of L spins with $\sigma_j = \pm 1$ for $j = 1, \ldots, L$. The second layer is a *hidden* layer. It is composed of N_h nodes, denoted by a vector $\vec{h} = (h_1, \ldots, h_{N_h})$ with $h_k = \pm 1$ for $k = 1, \ldots, N_h$. The h_k's are introduced as auxiliary particles in the probability model; they play roles similar to those of virtual particles in the valence-bond picture for MPSs [32, 57].

Given a specific spin configuration $\vec{\sigma}$, the RBM outputs the corresponding wave-function amplitude

$$\psi(\vec{\sigma}) = 2^{-N_h} \sum_{\{\vec{h}: h_k = \pm 1\}} \exp\Big(\sum_{j=1}^{L} a_j \sigma_j + \sum_{k=1}^{N_h} b_k h_k + \sum_{1 \le j \le L, 1 \le k \le N_h, j,k \in \mathbb{N}} W_{j,k} \sigma_j h_k \Big) \tag{2.1}$$

$$= \prod_{j=1}^{L} e^{a_j \sigma_j} \prod_{k=1}^{N_h} \cosh(b_k + \sum_{j=1}^{L} \sigma_j W_{j,k}). \tag{2.2}$$

Here, a_j and b_k are the bias parameters for the j-th spin and k-th hidden node, respectively, $W_{j,k}$ is a weight parameter describing the interlayer interaction between the j-th spin and k-th hidden node, and $\mathbb{N}$ denotes the set of all natural numbers. The a_j, b_k and $W_{j,k}$ are complex numbers. All such amplitudes defined on the computational basis yield a quantum state vector $|\Psi\rangle = \sum_{\vec{\sigma}} \psi(\vec{\sigma}) |\vec{\sigma}\rangle$, where the summation is over all 2^L spin configurations. It is remarkable that we adopt the RBM form with a factor of 2^{-N_h}. This choice allows us to use infinitely many hidden nodes h_k as long as b_k and $W_{j,k}$ decay sufficiently fast to ensure the convergence of $\psi(\vec{\sigma})$ as $N_h \to \infty$ for fixed system sizes L. In other words, it ensures that adding hidden nodes with associated parameters (b_k and $W_{j,k}$) being zero will not change the value of the wave function. This choice will facilitate the asymptotic analysis as shown below.

As mentioned in Sec. 2.1, the RBMs solved by relevant ML algorithms to approximate target states often feature a long-range form and a fast parameter decay. As more hidden nodes are added to the network, the RBM can capture higher-order correlations between spins [6], thus leading to higher accuracy in approximation. The parameter decay is manifested by the decay of weight parameters $W_{j,k}$ with an increasing index separation $|j-k|$ as well as the decay of b_k with increasing k.

In this work, we assume N_h to be an integer multiple of L which will facilitate the scaling analysis, especially for translationally invariant systems. When N_h is not an integer multiple of L, we can simply fill the last fragment with hidden nodes associated with zero-value parameters without influencing the wave-function values. We divide the hidden layer into multiple *levels*, each of which contains L hidden nodes (Fig. 2.1(a)). Thus, there are totally N_h/L levels while the ratio N_h/L is called the hidden-unit density in some references [6]. We will show that hidden nodes at the same level can capture the correlation of the same order between spins by performing an algorithm to reorder all hidden nodes for general RBMs. This point will be further clarified when we use the RBM form with translational symmetry to represent the ground states of 1D translationally invariant quantum systems as shown below [6].

One example of the quantum states that can be exactly represented by short-range RBMs [30, 31] is the 1D symmetry-protected topological (SPT) cluster state. The Hamiltonian of the SPT cluster system is defined on a 1D L-site lattice with periodic boundary conditions as $\hat{H}_{\text{cluster}} = -\sum_{j=1}^{L} \hat{\sigma}_{j-1}^z \hat{\sigma}_j^x \hat{\sigma}_{j+1}^z$, where $\hat{\sigma}^x$ and $\hat{\sigma}^z$ are Pauli matrices. A conventional r_0-range RBM is defined as an RBM satisfying $W_{j,k} = 0$ for any $|j-k| > r_0$. A short-range RBM usually refers to an r_0-range RBM with r_0 being a small constant independent of the

system size L. It was shown in Ref. [30, 31] that the ground state of $\hat{H}_{\text{cluster}}$ can be exactly represented by a 1-range RBM with L hidden nodes defined as:

$$
\begin{aligned}
&a_j = 0 \text{ (for any } j \in \{1, 2, \ldots, L\}), \\
&b_k = i\pi/4, W_{k-1,k} = i\pi/2, W_{k,k} = 3i\pi/4, \\
&W_{k+1,k} = i\pi/4 \text{ (for any } k \in \{1, 2, \ldots, L\}), \\
&W_{j,k} = 0 \text{ (for } |j-k| > 1),
\end{aligned} \tag{2.3}
$$

by using the stabilizer nature of the system to decrease the number of equation constraints for parameters from exponential to linear in L. Using our language of levels, this RBM just has one level and its weight parameters at this single level have a support of very short length which is a manifestation of its quantum entanglement satisfying an area law. Moreover, the translational symmetry of the system is inherited by the RBM form. The parameter patterns of this RBM also have a translational symmetry, which means that its parameters for different hidden nodes can be generated by the action of a translational-symmetry transformation operator on those for a single hidden node [6].

Inspired by the extensibility of the system of equations (2.3) with growing system sizes and considering the need to capture higher-order correlations between spins [6] and stronger quantum entanglement between subsystem blocks [31], we expect that the RBM representation of general quantum states has multiple, possibly infinitely many, levels and the length of the support of weight parameters at each level may increase from a small constant to the maximum length L. This motivates us to analyze generic long-range RBMs with properly specified nonlocal interactions between spins and hidden nodes (virtual particles).

2.2.1 Long-range-fast-decay RBMs

We now discuss aspects of the nonlocal structure of LRFD RBMs that were summarized at the end of Sec. 2.1. This leads to specific definitions of the two conditions that were mentioned there.

We begin by generalizing the RBM wave-function ansatz to an infinitely-many-hidden-node regime. An RBM state $|\Psi^{(L,\infty)}\rangle$ with infinitely many hidden nodes and a system size L can be defined as

$$|\Psi^{(L,\infty)}\rangle = \sum_{\vec{\sigma}} \psi^{(L,\infty)}(\vec{\sigma})|\vec{\sigma}\rangle, \tag{2.4}$$

where

$$\psi^{(L,\infty)}(\vec{\sigma}) = \prod_{j=1}^{L} e^{a_j^{(L)}\sigma_j} \prod_{k=1}^{\infty} \cosh(b_k^{(L)} + \sum_{j=1}^{L} \sigma_j W_{j,k}^{(L)}). \tag{2.5}$$

Its corresponding truncated-RBM sequence is defined as $\{|\Psi^{(L,N_h)}\rangle\}$, where

$$|\Psi^{(L,N_h)}\rangle = \sum_{\vec{\sigma}} \psi^{(L,N_h)}(\vec{\sigma})|\vec{\sigma}\rangle \tag{2.6}$$

and

$$\psi^{(L,N_h)}(\vec{\sigma}) = \prod_{j=1}^{L} e^{a_j^{(L)}\sigma_j} \prod_{k=1}^{N_h} \cosh(b_k^{(L)} + \sum_{j=1}^{L} \sigma_j W_{j,k}^{(L)}) \tag{2.7}$$

is constructed by removing the hyperbolic cosine terms with $k \geq N_h + 1$ from $\psi^{(L,\infty)}(\vec{\sigma})$.

Then, we define a subset of generic RBM states with infinitely many hidden nodes—long-range-fast-decay (LRFD) RBM states—as the RBMs whose parameters satisfy the following two conditions.

Condition 1 (*boundedness of* $W_{j,k}$)**.** There exists an L-independent integer $\tilde{k}_s \in \mathbb{N}$ and three nonnegative monotonically decreasing real functions $\lambda_R(\tilde{k})$, $\lambda_I(\tilde{k})$ and $\mu(r)$ such that, after a reordering of all hidden nodes, for all $k > \tilde{k}_s L$,

$$|\operatorname{Re}(W_{j,k}^{(L)})| \leq \lambda_R(\tilde{k})\mu(|j - j_c|_{\text{circ}}), \tag{2.8}$$

$$|\operatorname{Im}(W_{j,k}^{(L)})| \leq \lambda_I(\tilde{k})\mu(|j - j_c|_{\text{circ}}), \tag{2.9}$$

where $\tilde{k} \in \{1, 2, \ldots, N_h/L\}$ designates the numerical index of levels; j_c, the center spin for the k-th hidden node, denotes the site index of the spin with which the interaction of the k-th hidden node reaches its maximum among all $j \in \{1, 2, \ldots, L\}$; $|m|_{\text{circ}} = \min\{m, L - m\}$ in accordance with the periodic boundary conditions; and $r \in \{0, 1, \ldots, (L-1)/2\}$ denotes the distance between j and j_c assuming L is odd without influencing the validity of the following asymptotic analysis. The functions $\lambda_R(\tilde{k})$, $\lambda_I(\tilde{k})$ and $\mu(r)$ satisfy the conditions

that there exist finite L-independent nonnegative constants P_0 and μ_0 such that

$$\sum_{\tilde{k}=\tilde{k}_s+1}^{\infty} \left(\lambda_R^2(\tilde{k}) + \beta_1^2 \lambda_I^2(\tilde{k})\right) = P_0 < \infty, \tag{2.10}$$

$$\mu(r) \leq \mu_0 < \infty \quad (\text{for all } r \geq 0), \tag{2.11}$$

where $\beta_1 = 3\sqrt{2\ln 2}/\pi$ is found in the convergence proof given in Appendix A.1.

We provide an interpretation of each new variable as follows. $\tilde{k} = \tilde{k}(k)$ and $j_c = j_c(k)$ are both functions of k and the correspondence between the pair $(\tilde{k}, j_c)$ and k is a bijective map. It means that every hidden node is associated with a unique pair and thus can be uniquely positioned in the RBM network after the reordering (Fig. 2.1(a)). The hidden nodes capturing the correlation of the same order between spins are grouped at the same level so that the new indices of all hidden nodes characterized by the pair $(\tilde{k}, j_c)$ actually manifest the level of correlations. This characterization can also facilitate a symmetry manifestation for quantum states holding translational symmetry. This reordering step is to solve the problem that ML algorithms with a stochastic nature are often unable to automatically group the hidden nodes according to level stratification and their site positions usually exhibit randomness.

Condition 2 (*boundedness of* b_k)**.** After the same reordering of all hidden nodes that is described in Condition 1, for all $k > \tilde{k}_s L$,

$$|\operatorname{Re}(b_k^{(L)})| \leq \lambda_R(\tilde{k})\mu(0), \tag{2.12}$$

$$|\operatorname{Im}(b_k^{(L)})| \leq \lambda_I(\tilde{k})\mu(0). \tag{2.13}$$

The definition of LRFD RBMs should be understood from the point of view of state manifolds [32, 58]. A state manifold for quantum many-body states usually refers to a subspace of the whole Hilbert space spanned by a parameterized wave-function family [58], thus is a set containing specific types of quantum states. So the manifold of LRFD RBMs can be defined as a space spanned by all parameterized wave functions, every one of which belongs to a quantum-state sequence associated with a varying system size and satisfying the above Condition 1 and 2. One LRFD-RBM state refers to an element in this manifold. So this definition is in the same spirit as the definition of MPSs with different scaling laws [32, 57].

Condition 1 gives an upper bound on the magnitude of RBM weight parameters and actually provides a description of the nonlocal interaction between spins and hidden nodes (virtual particles). It requires that $|\mathrm{Re}(W_{j,k}^{(L)})|$ and $|\mathrm{Im}(W_{j,k}^{(L)})|$ are upper bounded, respectively, by the products $\lambda_R(\tilde{k})\mu(r)$ and $\lambda_I(\tilde{k})\mu(r)$. The monotonically decreasing functions $\lambda_R(\tilde{k})$ and $\lambda_I(\tilde{k})$ can be regarded as level-decay factors, while $\mu(r)$ is a factor describing the decay due to the increase of the distance between the spin-site index (j) and the corresponding spin-site index of the center spin ($j_c(k)$) for the k-th hidden node. The function $\mu(r)$ has a localization feature and resembles a single-modal localized orbital in the physics of periodic potentials, such as Wannier modes [59], which can be reflected by its monotonically decreasing with increasing r. So this description can effectively capture the parameter decays induced by both the level increase and the growth of system size, providing two degrees of freedom in characterizing the nonlocal interaction pattern. The separate treatments for the real and imaginary parts originate from their inequivalent positions in the RBM wave-function form, which is shown in Appendix A.1.

Condition 2 implies that the contribution of b_k-related terms can be upper bounded by the largest $W_{j,k}$-related terms at each level so that the $W_{j,k}$ weight parameters play a dominant role in the asymptotic analysis (Appendix A.1). Since there is often a degree of freedom in choosing the value of $\mu(0)$, Condition 2 can be satisfied for a wide range of RBM states.

Conditions 1 and 2 are proposed to ensure the convergence of the state vector (Eq. (2.5)) and provide a clear quantification for the rate of parameter decays, on the basis of which a complexity analysis can be conducted. A rigorous proof of the convergence of the state vector when Conditions 1 and 2 are satisfied is given in Appendix A.1. This proof is important not only because it ensures that the generalization of RBMs to an infinitely-many-hidden-node regime makes sense by defining them as the limits of some infinite sequences, but also because it introduces the key mathematical tricks and concepts that are necessary for analyzing the effects of truncations.

The core idea of the proof is that we can prove the sequence $\{\psi^{(L,nL)}(\vec{\sigma}) : n \in \mathbb{N}\}$ is a Cauchy sequence in the field of complex numbers $\mathbb{C}$ [60]. This proof is inspired by the fact that, when $b_k^{(L)}$ and $W_{j,k}^{(L)}$ decay sufficiently fast, the complex-valued ratio $\psi^{(L,(n+m)L)}(\vec{\sigma})/\psi^{(L,nL)}(\vec{\sigma})$ will quickly fall into the neighborhood of the point $z = 1$ in the complex plane as n increases. So we derive an upper and lower bound on the ratio's modulus $|\psi^{(L,(n+m)L)}(\vec{\sigma})/\psi^{(L,nL)}(\vec{\sigma})|$ which converge to 1 and an upper bound on the magnitude of its argument $|\arg\left(\psi^{(L,(n+m)L)}(\vec{\sigma})/\psi^{(L,nL)}(\vec{\sigma})\right)|$ which converges to 0 as n increases. Then we show that the corresponding magnitude sequence $\{|\psi^{(L,nL)}(\vec{\sigma})|\}$ and the argument sequence $\{\arg(\psi^{(L,nL)}(\vec{\sigma}))\}$ are Cauchy sequences in the field of real numbers $\mathbb{R}$, thus $\{\psi^{(L,nL)}(\vec{\sigma})\}$ is a Cauchy sequence in $\mathbb{C}$.

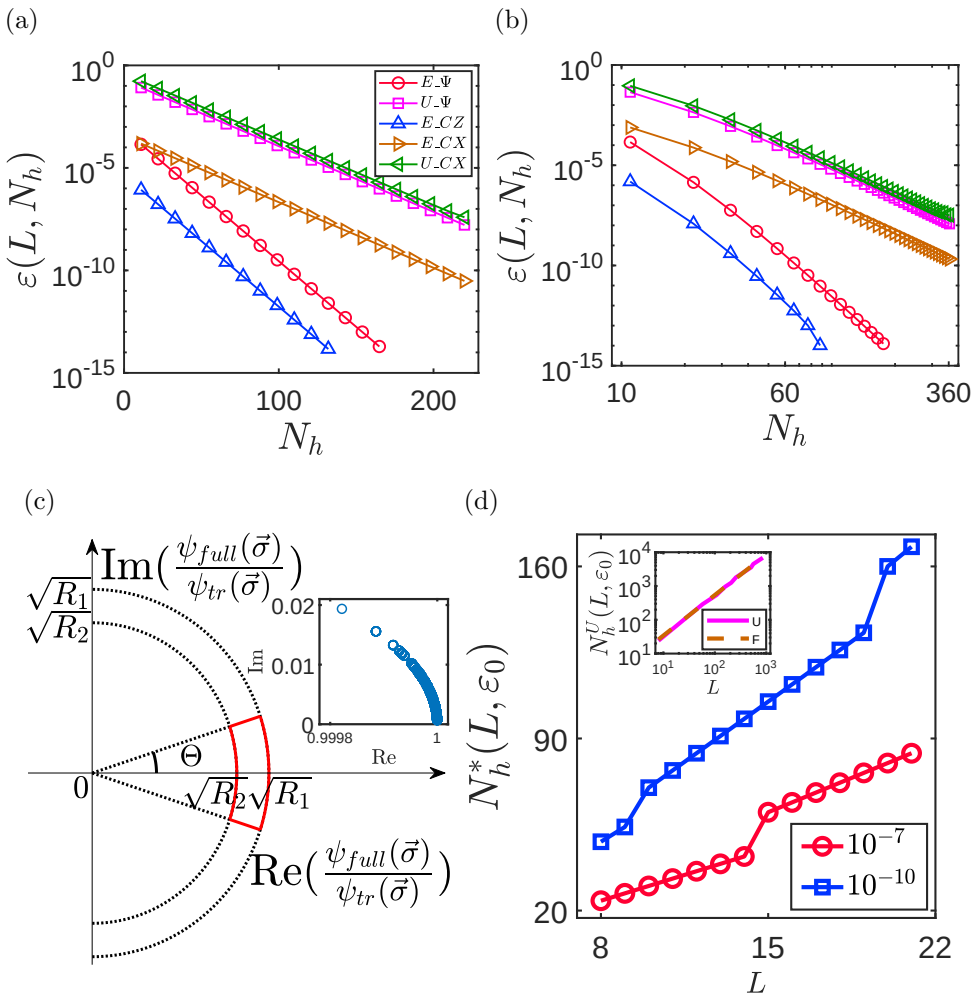


Figure 2.2: (a)(b) Comparison between the exact and estimated truncation errors $\varepsilon(L, N_h)$ as a function of N_h with fixed L for 1D SPT cluster states with a perturbation part. E_Ψ: exact values, first-type truncation errors; U_Ψ: upper-bound-based estimation, first-type; E_CZ: exact, second-type, $\hat{B} = \hat{\sigma}_1^z\hat{\sigma}_2^z$; E_CX: exact, second-type, $\hat{B} = \hat{\sigma}_1^x\hat{\sigma}_2^x$; U_CX: upper-bound-based estimation, second-type, $\hat{B} = \hat{\sigma}_1^x\hat{\sigma}_2^x$. The perturbation part is constructed as Eqs. (2.24)–(2.26) show. $\mu(r) = \frac{1}{2}\delta_Q r^{-\alpha_Q}$ for $r \neq 0$, $\mu(0) = \delta_Q = 0.1$, $\alpha_Q = 3$, $c_w = c_b = 1 + i$, $a_0 = 0$ and $L = 11$. (a) Exponential decay of $\lambda(\tilde{k})$ (vertical axis: log scale). $\lambda(\tilde{k}) = 0.2\delta_P^{-(\tilde{k}-1)}$ with $\delta_P = 1.5$. (b) Power-law decay of $\lambda(\tilde{k})$ (on a log-log scale). $\lambda(\tilde{k}) = \tilde{k}^{-\alpha_P}$ with $\alpha_P = 3$. (c) Schematic interpretation of variables used in proofs. The inset in (c) shows the distribution of data points of the ratio $\psi^{(L,\infty)}(\vec{\sigma})/\psi^{(L,N_h)}(\vec{\sigma})$ with $N_h = L$, which corresponds to a truncation removing hidden nodes starting from the second level. The parameter setting is the same as that in (b). The data points are all localized in the neighborhood of $z = 1$ in the complex plane enclosed by the red solid curves in (c) as we analyzed. (d) Scaling of $N_h^*(L, \varepsilon_0)$ in L for two fixed values of the first-type truncation errors ε_0 with the same parameter setting as in (b). Red circle: $\varepsilon_0 = 10^{-7}$; Blue square: $\varepsilon_0 = 10^{-10}$. The inset in (d) shows the scaling of N_h estimated based on our upper bounds with $\varepsilon_0 = 10^{-3}$. Magenta solid: using exact values of upper bounds; Brown dashed: using leading-order estimations. The two curves almost coincide.

2.2.2 Effects of wave-function truncation for fixed system sizes

We derive upper bounds on truncation errors associated with two measures of state differences for the sequence of truncated LRFD RBM states. Define $\varepsilon(L, N_h)$ to be a specific type of truncation error for using $|\Psi^{(L,N_h)}\rangle$ to approximate $|\Psi^{(L,\infty)}\rangle$.

A natural measure of state differences is the square of the l_2-norm [61] of the state-vector difference, $\||\tilde{\Psi}^{(L,\infty)}\rangle - |\tilde{\Psi}^{(L,N_h)}\rangle\|_2^2$, where the tilde symbol is used to represent corresponding states after a normalization operation. It is remarkable that the RBM wave-function ansatz is not automatically normalized and an estimation of the normalization factor $\langle\Psi|\Psi\rangle$ is important and often tricky as shown in Appendix A.2. This measure of truncation errors is adopted in fundamental works about the faithfulness and efficiency of other wave-function ansätze, such as MPSs [62, 63]. So it allows us to make a direct comparison between the efficiencies of RBMs and other state representations.

A second measure of state differences is a Hermitian-operator-based expectation-value difference defined as

$$\begin{aligned} &|\langle\hat{B}\rangle^{(L,\infty)} - \langle\hat{B}\rangle^{(L,N_h)}| \\ &= |\langle\tilde{\Psi}^{(L,\infty)}|\hat{B}|\tilde{\Psi}^{(L,\infty)}\rangle - \langle\tilde{\Psi}^{(L,N_h)}|\hat{B}|\tilde{\Psi}^{(L,N_h)}\rangle|. \end{aligned} \tag{2.14}$$

Here, $\hat{B}$ can be any Hermitian operator of the form $\hat{B} = \bigotimes_{j=1}^{L} \hat{\sigma}_j^{(m_j)}$, where $\bigotimes$ is the tensor product symbol, $m_j \in \{0, 1, 2, 3\}$, $\hat{\sigma}_j^{(0)} = I_{2\times 2}$ is a 2-by-2 identity matrix, and $\{\hat{\sigma}_j^{(1)}, \hat{\sigma}_j^{(2)}, \hat{\sigma}_j^{(3)}\}$ denote the Pauli matrices. We also use this measure as $\{\hat{\sigma}_j^{(m)} : m = 0, 1, 2, 3\}$ is a complete basis set for the local Hilbert space for the j-th spin and a wide

range of typical physical observables, such as spin correlations and total energy, correspond to Hermitian operators of such type or linear combinations of polynomially many such operators.

Then we can prove a lemma which provides upper bounds on truncation errors of the above two types for the sequence of truncated LRFD RBM states.

Lemma 3 (*upper bounds on truncation errors*)**.** For LRFD RBMs satisfying Conditions 1 and 2, after the same reordering of all hidden nodes described in Condition 1, for all $N_h > \tilde{k}_s L$,

$$\| |\tilde{\Psi}^{(L,\infty)}\rangle - |\tilde{\Psi}^{(L,N_h)}\rangle \|_2^2 \leq F_1\Big(LQ(L)P(N_h/L)\Big), \tag{2.15}$$

$$|\langle \hat{B}\rangle^{(L,\infty)} - \langle \hat{B}\rangle^{(L,N_h)}| \leq F_2\Big(LQ(L)P(N_h/L)\Big), \tag{2.16}$$

where

$$F_1(x) = 2 - 2\exp[-2(1+\beta_1^2)x]\cos(4\beta_2 x) \tag{2.17}$$

$$= c_1 x + \mathcal{O}(x^2) \quad (\text{as } x \to 0), \tag{2.18}$$

$$\begin{aligned} F_2(x) = {} & \max\{|\exp(4x) - 1|, |1 - \exp(-4\beta_1^2 x)|\} \\ & + \max\{\Big[\exp(8x) - 2\exp(4x)\cos(8\beta_2 x) + 1\Big]^{1/2}, \\ & \Big[\exp(-8\beta_1^2 x) - 2\exp(-4\beta_1^2 x)\cos(8\beta_2 x) + 1\Big]^{1/2}\} \end{aligned} \tag{2.19}$$

$$= c_2 x + \mathcal{O}(x^{3/2}) \quad (\text{as } x \to 0). \tag{2.20}$$

$$P(m) = \sum_{\tilde{k}=m+1}^{\infty} \lambda^2(\tilde{k}) \quad (m \geq \tilde{k}_s, m \in \mathbb{N}), \tag{2.21}$$

$$Q(L) = \Big(\sum_{r=0}^{(L-1)/2} \mu(r)\Big)^2, \tag{2.22}$$

the relevant constants are $\beta_2 = 3\sqrt{3}/\pi$, $c_1 = 4(1+\beta_1^2)$ and $c_2 = 4\beta_1^2 + 4(\beta_1^4 + 4\beta_2^2)^{1/2}$ and we have assumed that $\lambda_R(\tilde{k}) = \lambda_I(\tilde{k}) = \lambda(\tilde{k})$ for simplicity which holds throughout the following discussion.

The proof is given in Appendix A.2 which uses arguments similar to those described in the proof for the convergence of LRFD RBMs. Based on the intuition that the ratio $\psi^{(L,\infty)}(\vec{\sigma})/\psi^{(L,N_h)}(\vec{\sigma})$ will fastly converge to 1 with increasing N_h, we derive an upper bound $\sqrt{R_1}$ and a lower bound $\sqrt{R_2}$ on the ratio's modulus $|\psi^{(L,\infty)}(\vec{\sigma})/\psi^{(L,N_h)}(\vec{\sigma})|$ and an upper bound Θ on the magnitude of its argument $|\arg\left(\psi^{(L,\infty)}(\vec{\sigma})/\psi^{(L,N_h)}(\vec{\sigma})\right)|$. The two types of truncation errors can be upper bounded using these three variables and the two

upper bounds can be finally expressed as functions ($F_1(x)$ and $F_2(x)$) of $LQ(L)P(N_h/L)$ which decreases to zero with increasing N_h and fixed L. The idea of the proof is shown schematically in Fig. 2.2(c). For simplicity, we assume that $\lambda_R(\tilde{k}) = \lambda_I(\tilde{k}) = \lambda(\tilde{k})$ throughout the following discussion.

Based on our description of the nonlocal interactions between spins and virtual particles and using the language of levels, $P(N_h/L)$ is a summation of all level-decay factors for hidden nodes at levels starting from $\tilde{k} = N_h/L + 1$ to $\tilde{k} = \infty$, while $Q(L)$ represents the localized "orbital" at every single level and contributes a factor reflecting the pure influence of system-size growing regardless of levels. The two different types of truncation errors correspond to two different forms of the function $F(x)$, but both of them are analytic at the point $x = 0$.

We give the scaling of truncation errors in N_h as below. It can be obtained that, if $Q(x) = \mathcal{O}(q(x))$ as $x \to \infty$, $P(x) = \mathcal{O}(1/p_d(x))$ as $x \to \infty$, and $F(x) = \mathcal{O}(f(x))$ as $x \to 0$, then

$$\varepsilon(L, N_h) = \mathcal{O}(f(\frac{L\, q(L)}{p_d(N_h/L)})) \quad (\text{as } N_h \to \infty). \tag{2.23}$$

Our construction of LRFD RBMs and theoretical analysis of the truncation errors can be further clarified with results from numerical computations. We can construct LRFD

RBMs with translation symmetry whose parameters exactly satisfy

$$W_{j,k}^{(L)} = c_w \lambda(\tilde{k}) \mu(|j - j_c|_{\text{circ}}), \tag{2.24}$$

$$b_k^{(L)} = c_b \lambda(\tilde{k}) \mu(0), \tag{2.25}$$

$$a_j^{(L)} = a_0 \tag{2.26}$$

for any $1 \leq j \leq L$, $1 \leq k \leq N_h$, where a_0, c_w and c_b are complex constants with $|c_b| \leq |c_w|$ to satisfy Condition 2, $\tilde{k} = \tilde{k}(k) = \lceil k/L \rceil$, $j_c = j_c(k) = k - (\tilde{k} - 1)L$, N_h is an integer multiple of L, $\lceil x \rceil$ denotes the ceiling function, and $\tilde{k}_s = 0$ in this case. It can be shown that such an RBM form can be directly transformed into the RBM form proposed to represent ground states of 1D translationally invariant systems [6] for any finite N_h but we generalize it to an infinitely-many-hidden-node regime. Since the parameters for different hidden nodes can be generated by the action of a translational-symmetry transformation operator on those for a single hidden node, we just need to focus on one representative hidden node for each level. So we propose an importance measure $\eta(j, \tilde{k}, L)$ to measure the importance of a set of edges which is defined as

$$\eta(j, \tilde{k}, L) = \left| \text{Re}(W_{j,(L+1)/2+(\tilde{k}-1)L}^{(L)}) \right|^2 + \beta_1^2 \left| \text{Im}(W_{j,(L+1)/2+(\tilde{k}-1)L}^{(L)}) \right|^2 \tag{2.27}$$

and present it as a function of the spin-site index j and level index $\tilde{k}$. Its 3D structure can reflect the decay of both $\lambda(\tilde{k})$ and $\mu(r)$ while the center of the "orbital" at every level is localized around $j = (L + 1)/2$. So a plotting of the peak at every level as a function of the level number $(\tilde{k})$ can reflect the decay of $\lambda(\tilde{k})$. One example of such LRFD RBM with

a power-law decaying $\lambda(\tilde{k})$ is shown in Fig. 2.1(b).

We show the two types of truncation errors $\varepsilon(L, N_h)$ as a function of N_h with fixed L for 1D SPT cluster states with a perturbation part (Fig. 2.2(a) and 2.2(b)). It means that the RBM is constructed as a summation of the setting defined in the system of equations (2.3) and a perturbation part specified as Eqs. (2.24)–(2.26) show. The numerical results for $\lambda(\tilde{k})$ with exponential and power-law decays are given. As described above, the 1D SPT cluster states can be exactly represented by a short-range (1-range) RBM [30]. Using our description, its RBM representation just has one level, and the corresponding $\lambda(\tilde{k})$ and $\mu(r)$ quickly go down to zero for $\tilde{k} > 1$ and $r > 1$. The addition of the perturbation part makes the composite RBM a LRFD RBM so that we can study the truncation errors. We give the results for both types of truncation errors and let $\hat{B}$ be the operator of spin correlations between spin 1 and 2 in z and x directions.

Our numerical experiments on the scaling of the truncation errors in N_h with fixed L are well upper bounded by our estimations given in inequalities (2.15) and (2.16), which substantiates our theoretical analysis. Those experiments also indicate that our estimations in Eq. (2.23) correctly capture the asymptotic properties of $\varepsilon(L, N_h)$ with varying N_h. Moreover, the fact that the curve of exact $\varepsilon(L, N_h)$ and that of our estimation associated with $\hat{B} = \hat{\sigma}_1^x \hat{\sigma}_2^x$ have exactly the same slope implies that our estimation in Eq. (2.23) gives an asymptotically optimal upper bound. It means that, for the second-type truncation errors (inequality (2.16)), there is still room to improve the constant prefactors in our estimation, but we cannot qualitatively further improve the upper bound. In comparison, there is room to both qualitatively improve the upper bound and improve the constant prefactors for the first-type truncation errors (inequality (2.15)).

2.2.3 Scaling of complexity

Table 2.1: Complexity estimations for distinct typical settings of $\mu(r)$ and $\lambda(\tilde{k})$. "$-$" in the $\mu(r)$ column denotes all $\mu(r)$ functions that make $Q(L)$ converge as $L \to \infty$. "$-$" in the $\lambda(\tilde{k})$ column denotes all $\lambda(\tilde{k})$ functions that make $P(N_h/L)$ have the asymptotic behavior of $\mathcal{O}(1/\ln(N_h/L))$ as $N_h \to \infty$. In all settings, $\delta_P > 1$, $\alpha_P > 1/2$ and each entry provides a description of the asymptotic behavior of the corresponding function.

Manifold	$\mu(r)$	$Q(L)$	$\lambda(\tilde{k})$	$P(N_h/L)$	$N_h^*(L,\varepsilon)$
$S_2^{(1)}$	-	converge	$\delta_P^{-\tilde{k}}$	$\mathcal{O}(\delta_P^{-2N_h/L})$	$\mathcal{O}(L\ln(L/\varepsilon))$
$S_2^{(2)}$	-	converge	$\tilde{k}^{-\alpha_P}$	$\mathcal{O}((N_h/L)^{1-2\alpha_P})$	$\mathcal{O}((L^{2\alpha_P}/\varepsilon)^{1/(2\alpha_P-1)})$
$S_2^{(3)}$	r^{-1}	$\mathcal{O}((\ln L)^2)$	$\delta_P^{-\tilde{k}}$	$\mathcal{O}(\delta_P^{-2N_h/L})$	$\mathcal{O}(L\ln(L/\varepsilon))$
$S_2^{(4)}$	r^{-1}	$\mathcal{O}((\ln L)^2)$	$\tilde{k}^{-\alpha_P}$	$\mathcal{O}((N_h/L)^{1-2\alpha_P})$	$\mathcal{O}((L^{2\alpha_P}(\ln L)^2/\varepsilon)^{1/(2\alpha_P-1)})$
$S_2^{(5)}$	$\to \mu_\infty > 0$	$\mathcal{O}(L^2)$	$\delta_P^{-\tilde{k}}$	$\mathcal{O}(\delta_P^{-2N_h/L})$	$\mathcal{O}(L\ln(L/\varepsilon))$
$S_2^{(6)}$	$\to \mu_\infty > 0$	$\mathcal{O}(L^2)$	$\tilde{k}^{-\alpha_P}$	$\mathcal{O}((N_h/L)^{1-2\alpha_P})$	$\mathcal{O}((L^{2\alpha_P+2}/\varepsilon)^{1/(2\alpha_P-1)})$
$S_2^{(7)}$	-	converge	-	$\mathcal{O}(1/\ln(N_h/L))$	$\mathcal{O}(L\exp(L/\varepsilon))$

We can investigate the scaling of spatial complexity in system sizes for LRFD RBMs as the results in Sec. 2.2.1 and Sec. 2.2.2 still hold for varying L. We give an upper-bound estimation of the complexity of RBM representations which depends on the asymptotic behavior at $x = \infty$ of the functions $P(x)$ (Eq. (2.21)) and $Q(x)$ (Eq. (2.22)), and thus is determined by the decaying rates specified by $\lambda(\tilde{k})$ and $\mu(r)$.

Define the minimum N_h to achieve a sufficiently small approximation error ε_0 as

$$N_h^*(L,\varepsilon_0) = \inf\{N_h : \varepsilon(L,N_h) \le \varepsilon_0\}. \tag{2.28}$$

Using Lemma 3, the sufficient condition for $\varepsilon(L,N_h) \le \varepsilon_0$ is that the corresponding upper bound on truncation errors is no larger than ε_0. So this provides one way to get an upper

bound on $N_h^*(L,\varepsilon_0)$ for LRFD RBMs. It can be shown that

$$N_h^*(L,\varepsilon_0) = \mathcal{O}(L\, p_d^{-1}(\frac{L\, q(L)}{f^{-1}(\varepsilon_0)})) \quad (\text{as } L \to \infty), \tag{2.29}$$

where $q(x)$, $p_d(x)$ and $f(x)$ are functions to specify the asymptotic behaviors of $Q(x)$, $P(x)$ and $F(x)$ as defined above and the superscript "-1" denotes the inverse of the corresponding function.

Rich information can be extracted from Eq. (2.29). First, the first factor L comes from our assumption that N_h is an integer multiple of the system size L and the second factor L in front of $q(L)$ is extracted using the translational symmetry of the wave function. So these two factors reflect the growing system sizes and the remaining factors reflect the distinction in complexity for different LRFD RBMs.

Second, $P(N_h/L)$ and $Q(L)$ (thus $\mu(r)$ and $\lambda(\tilde{k})$) which characterize the nonlocal structure of RBMs in our description have qualitatively different influence on the complexity. Specifically, $Q(L)$ can converge to a finite L-independent constant in the thermodynamic limit and does not influence the complexity for sufficiently localized "orbitals" in the cases where $\mu(r)$ decays sufficiently fast. With the upper boundedness condition for $\mu(r)$ (Eq. (2.11)), $Q(L)$ can contribute an at-most-linear factor to this upper bound on $N_h^*(L,\varepsilon_0)$. By contrast, the asymptotic property of $P(x)$ significantly influences the complexity and may lead to the inefficiency of RBM representations if $\lambda(\tilde{k})$ decays sufficiently slowly. That would imply that there are too many high-order correlations between spins to be captured by the RBM so polynomially many parameters are not enough to fully compress the information into the RBM form. But as long as $p_d^{-1}(x)$ has an at-most-power-law

dependence on x, this upper-bound estimation will imply that the complexity is definitely at most polynomial in both system size L and $1/\varepsilon_0$ with the above two types of truncation errors. Moreover, it is also remarkable that our estimation only provides an upper bound on the complexity, so a faster-than-polynomial scaling of the bound (such as $S_2^{(7)}$ in Table 2.1) does not necessarily imply the inefficiency of the representation. It is possible that the upper bound is not tight and the real complexity is at most polynomial in this case.

Third, the asymptotic behavior of $F(x)$ at $x = 0$ also influences the scaling of the complexity and it directly acts on ε_0. We have demonstrated that, for the two types of truncation errors described above, the corresponding $F(x)$'s ($F_1(x)$ and $F_2(x)$) are both analytic at $x = 0$. For general types of truncation errors that can be upper bounded by a function $F(LQ(L)P(N_h/L))$, $1/f^{-1}(\varepsilon_0)$ has a power-law dependence on $1/\varepsilon_0$ as long as $F(x)$ is analytic at $x = 0$ based on the Taylor series expansion of the function.

This result suggests separate effects of the factors $\mu(r)$ and $\lambda(\tilde{k})$. The scaling of entanglement entropy, which is an important measure of the complexity of quantum many-body states, is influenced by $\mu(r)$, whereas $\lambda(\tilde{k})$ significantly influences the spatial complexity of parameterization in LRFD RBM representations. The length of the support of $\mu(r)$, which determines the "range" r_0 of RBMs, directly influences the scaling of the entanglement entropy of the states between subregions [31] but does not directly contribute a faster-than-polynomial factor to the parameterization complexity. This result possibly provides further theoretical evidence for the high efficiency of RBMs in representing states with entanglement entropy scaling faster than an area law in system sizes [31].

We apply our complexity estimation to several typical settings of $\mu(r)$ and $\lambda(\tilde{k})$ in Table 2.1. The manifolds $S_2^{(j)}$ with $1 \leq j \leq 6$ all correspond to a spatial complexity which is

at most polynomial in L. We also apply this analysis to RBMs constructed as the 1D SPT cluster states with a perturbation part. Our numerical results on the scaling of $N_h^*(L, \varepsilon_0)$ in L with fixed ε_0 (Fig. 2.2(d)) for small system sizes are consistent with our theoretical analysis summarized in Table 2.1. The piecewise linearity of $N_h^*(L, \varepsilon_0)$ as a function of L with a slope growing very slowly implies that the scaling is perhaps just slightly faster than linear, consistent with our estimation based on parameter settings. The piecewise linearity is due to our assumption that N_h is an integer multiple of L. So it applies a ceiling operation to the ratio N_h/L which will not change when L varies within a small range. The inset in Fig. 2.2(d) shows that $N_h^U(L, \varepsilon_0)$ as upper bounds on $N_h^*(L, \varepsilon_0)$ in our analysis obtained by using the exact values of the right-hand side of inequality (2.15) and its leading-order estimations are almost the same and both have a power-law scaling in L as indicated by Eq. (2.29), which support the validity of our complexity analysis.

2.2.4 Spin-correlation information

In this subsection, we analyze what information about the physical properties of the quantum states can be extracted from the LRFD RBM form using our description of the nonlocal structure. Here, we focus on a small-parameter regime in which a_j, b_k, $W_{j,k} \leq \varepsilon_1$, and $\varepsilon_1 \ll 1/L$, $\varepsilon_1 \ll 1/N_h$. We do not explicitly write the superscript "(L)" for RBM parameters and assume that the RBM just has a finite number (N_h) of hidden nodes in this subsection.

Based on the proof given in Appendix A.3, we find that the correlation in the z direction between spins with a distance of r for a LRFD RBM with translational symmetry

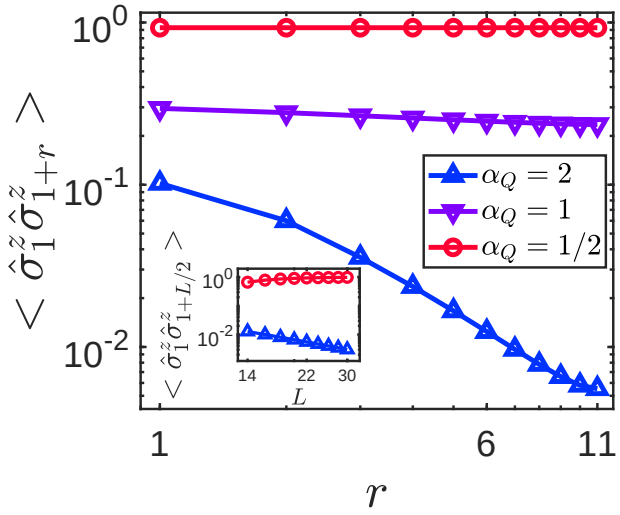


Figure 2.3: Spin correlations in the z direction as a function of distance r on a log-log scale. The LRFD RBMs are constructed as Eqs. (2.24)–(2.26) show, where $\mu(r) = \frac{1}{2}\delta_Q r^{-\alpha_Q}$ for $r \neq 0$, $\mu(0) = \delta_Q = 0.2$, $\lambda(\tilde{k}) = \tilde{k}^{-\alpha_P}$, $\alpha_P = 3.5$, $c_w = 1$, $c_b = 0$, $a_0 = 0$, $L = 22$ and $N_h = 5L$. The inset shows the spin correlation $\langle \hat{\sigma}_1^z \hat{\sigma}_{1+L/2}^z \rangle$ with r being the half-chain length for varying L (on a log-log scale). It shows a convergence of $\langle \hat{\sigma}_1^z \hat{\sigma}_{1+L/2}^z \rangle$ to an L-independent constant (almost attaining the maximum value 1) for $\alpha_Q = 1/2$ and a decay for $\alpha_Q = 2$.

is

$$C^z_{\text{unnorm}}(r) = \langle \Psi^{(L,N_h)} | \hat{\sigma}_1^z \hat{\sigma}_{1+r}^z | \Psi^{(L,N_h)} \rangle \tag{2.30}$$

$$= 2\big(\operatorname{Re}(WW^T)\big)_{1,1+r} + 4\operatorname{Re}(a_1)\operatorname{Re}(a_{1+r}) + \mathcal{O}(\varepsilon_1^3) \quad (\text{as } \varepsilon_1 \to 0). \tag{2.31}$$

Note that the above result is the r-related part of the spin correlation, while its real value is $C^z_{\text{unnorm}}(r)$ divided by an r-independent normalization factor $\langle \Psi^{(L,N_h)} | \Psi^{(L,N_h)} \rangle$. So for RBMs constructed as Eqs. (2.24)–(2.26) show with $a_0 = 0$ for simplicity,

$$C^z_{\text{unnorm}}(r) \approx |c_w|^2 \sum_{\tilde{k}=1}^{N_h/L} |\lambda(\tilde{k})|^2 \sum_{j_c=1}^{L} \mu(|1 - j_c|_{\text{circ}}) \mu(|1 + r - j_c|_{\text{circ}}). \tag{2.32}$$

So the $\mu(r)$-related factor as shown above describes the decaying rate of spin correlations in the z direction as a function of the distance r, while the $\lambda(\tilde{k})$-related factors independent of r do not influence the decaying rate if we only consider the leading-order terms in Eq. (2.31).

The above result in Eq. (2.31) gives an interpretation of the roles of hidden nodes. The hidden nodes can be viewed as intermediate virtual particles that relate spins (physical particles) at different lattice sites. When an RBM is short-range, the term $\left(\mathrm{Re}(WW^T)\right)_{1,1+r}$ will vanish for large enough r as there is no virtual particle that can have both nonzero connectivity to two spins separated by r. Then, more intermediate hidden nodes are needed to transport such relations, which means that we need to consider higher-order terms. This is additional evidence that long-range RBMs can represent states with strong quantum correlations. It is shown in Appendix A.3 that, even when $\mu(r) \to 0$ as $r \to 0$, we can still construct LRFD RBMs in which the spin correlations in the z direction can have long-range decayings lower bounded by $\Theta(1/r^{\alpha_Q})$ (for $\mu(r) = \Theta(1/r^{\alpha_Q})$) with $\alpha_Q > 1$, $\Theta(\ln r/r)$ (for $\mu(r) = \Theta(1/r)$), and even $\Theta(1)$ (for $\mu(r) = \Theta(1/r^{\alpha_Q})$ with $0 < \alpha_Q \leq \frac{1}{2}$). These three kinds of decaying rates of spin correlations are demonstrated by numerical computations (Fig. 2.3). The spin correlation $\langle \hat{\sigma}_1^z \hat{\sigma}_{1+r}^z \rangle$ almost saturates the maximum value 1 for $\alpha_Q = 1/2$ and have different long-range decaying rates for $\alpha_Q = 1$ and $\alpha_Q = 2$ as r increases.

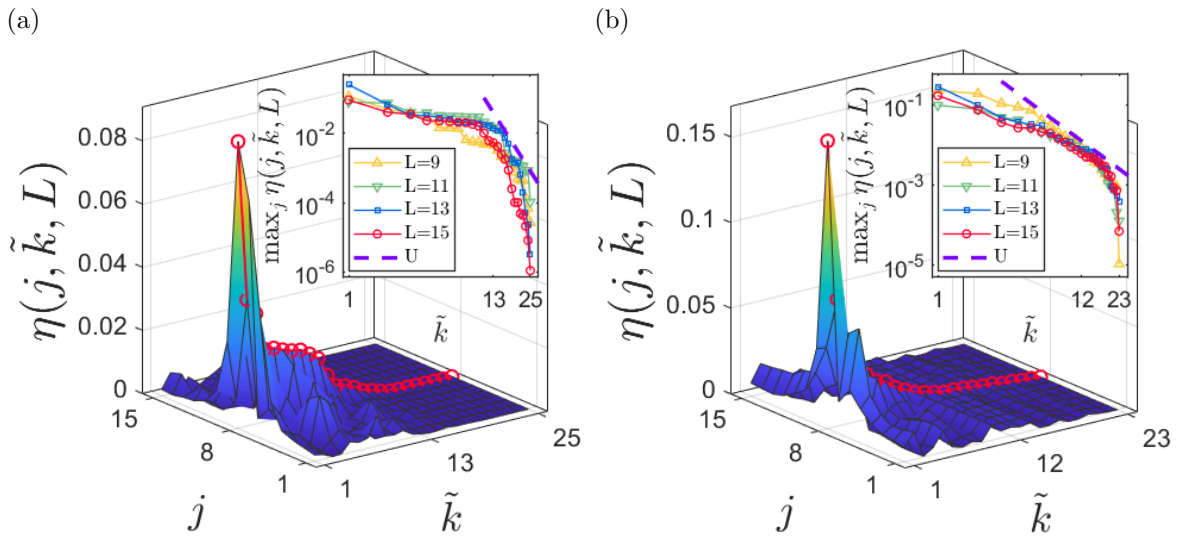


Figure 2.4: Importance measure $\eta(j,\tilde{k},L)$ for the RBMs approximating ground states of two critical systems with $L=15$. (a) TFIM with $B_x=1$. (b) XXZ model with $J_z=-0.2$. The insets in each subfigure show the decays of the maximum importance measure at each level as level number $\tilde{k}$ increases on a log-log scale. The system size $L=9$, 11, 13 and 15. The purple dashed curve implies that these decaying curves can be upper bounded by a power-law decay. By numerical fitting, the corresponding α_P for (a) and (b) are 2.957 and 1.232, respectively.

2.3 Ground-state applications

Based on the proposal of the concept of LRFD RBMs and the theoretical analysis of their spatial complexity, it is natural to explore their applications to learning quantum states associated with specific models. First, we theoretically prove that the state with all spins pointing up in the z direction, which is the ground state of a spin-$\frac{1}{2}$ system with a single magnetic field in the z direction and has a form of the Kronecker delta function, can be approximated by LRFD RBMs with arbitrary accuracy in Appendix A.4. We find that the RBM construction is not unique for such a target state even when fixing the global phase which implies eliminating the degree of freedom associated with a global gauge transformation and we give one construction. Thus, we provide one example of the utility

of LRFD RBMs in state representation for arbitrarily large system sizes.

Second, we are particularly interested in the behavior of RBMs in cases where other state representations become less efficient. We numerically study the representation of the ground states of critical systems with finite sizes for which the MPS representation becomes less efficient [62, 63], while MPS has achieved notable success in representing quantum many-body states with entanglement entropy satisfying an area law [32, 57].

We use RBMs with translational symmetry and apply the conventional quantum Monte Carlo algorithm (also a variational method) with stochastic-reconfiguration optimizations [6, 64, 65, 66] to learn the ground states of two typical quantum models: the 1D transverse-field Ising model (TFIM) (Eq. (2.33)) and XXZ model (Eq. (2.34)), described by Hamiltonians

$$\hat{H} = -\sum_{1\leq j\leq L} \hat{\sigma}_j^z \hat{\sigma}_{j+1}^z - B_x \sum_{j=1}^{L} \hat{\sigma}_j^x, \tag{2.33}$$

and

$$\hat{H} = \sum_{1\leq j\leq L} (-\hat{\sigma}_j^x \hat{\sigma}_{j+1}^x - \hat{\sigma}_j^y \hat{\sigma}_{j+1}^y + J_z \hat{\sigma}_j^z \hat{\sigma}_{j+1}^z) \tag{2.34}$$

with periodic boundary conditions, respectively, where B_x denotes the strength of a transverse field and J_z denotes the strength of coupling in the z direction. We use RBMs to learn the ground state of the TFIM with $B_x = 1$ which implies that the quantum system is exactly in the phase-transition point between a ferromagnetic and a paramagnetic phase [67] and of the XXZ model with $J_z = -0.2$ which implies that the system is in a gapless disordered

XY phase [68]. Both systems are critical systems with the entanglement entropy of the ground states scaling logarithmically in system sizes [67, 69, 70]. The ground states of these two Hamiltonians (at least for small system sizes) can be well learned by RBMs, which is demonstrated by the high accuracy in spin-correlation calculations given in Appendix A.5. The importance measures $\eta(j, \tilde{k}, L)$ for these two RBMs are provided in Fig. 2.4(a) and 2.4(b).

The numerical results show that the RBM representations of the two ground states of the above two critical systems have forms very similar to LRFD RBMs. The overall 3D structures for the importance measures $\eta(j, \tilde{k}, L)$ are similar to the one presented in Fig. 2.1(b) which corresponds to a standard LRFD RBM. The weight parameters for hidden nodes at the same level are quite localized and decay fastly as the level number $\tilde{k}$ increases and as the spin site index j goes away from the center. Moreover, it seems that the "ridge" of $\eta(j, \tilde{k}, L)$ for varying system sizes can be upper bounded by an L-independent power-law decay curve, based on which we can extract a corresponding α_P characterizing the rate of level decay for these small-system-size wave functions. If these features still hold as L increases and approaches infinity, these states will form LRFD RBMs which belong to the set $S_2^{(2)}$ or $S_2^{(6)}$ in Table 2.1 and the corresponding $\lambda(\tilde{k})$ and $\mu(r)$ can be defined.

Moreover, the above results exhibit a feature that is also manifested in the theory of MPS representations. It has been shown that [62], though MPS becomes less efficient in representing the ground states of critical systems, the bond dimension required to achieve an approximation error ε_0 can still be upper bounded by a function scaling polynomially in the system size L. The exponent in the power-law dependence of spatial complexity of MPSs on L depends on the central charge c, which is a quantity roughly quantifying the "degrees of

freedom of the theory" in conformal field theory [57]. A larger c leads to a higher exponent in that estimation which implies a higher complexity in MPS representation. While the TFIM at the above phase-transition point has $c = \frac{1}{2}$ and the XXZ model in the disordered XY phase has $c = 1$ [71], our numerical results do show a smaller fitted α_P for the XXZ model, which implies that the XXZ model has more intrinsic "complexity" compared to the TFIM, thus needing more parameters to capture this complexity.

2.4 State manifolds and complexity classification

Rigorously speaking, the numerical results for systems of finite sizes only provide evidence supporting that the states may be LRFD RBMs but cannot prove it, since the properties of RBMs in the process of approaching the thermodynamic limit are not yet known. Based on the success of RBMs in numerical simulations and the fact that they can often achieve high accuracy even with a constant number of levels (at least for small system sizes), we conjecture that the ground states of a wide range of quantum systems may be exactly represented by LRFD RBMs or a variant of them. Here, the term "variant" means generalizing the forms specified in Condition 1 and 2 by adding more factors that can be naturally incorporated into our complexity analysis. For example, the $\lambda(\tilde{k})$ and $\mu(r)$ functional forms, which are L-independent in our definition of LRFD RBMs, can be generalized into $\lambda(\tilde{k}, L)$ and $\mu(r, L)$, respectively, while their effects can be easily evaluated using our paradigm for complexity analysis.

We summarize the relations between multiple typical state manifolds so that the significance of proposing the concept of LRFD RBMs can be better understood. A state

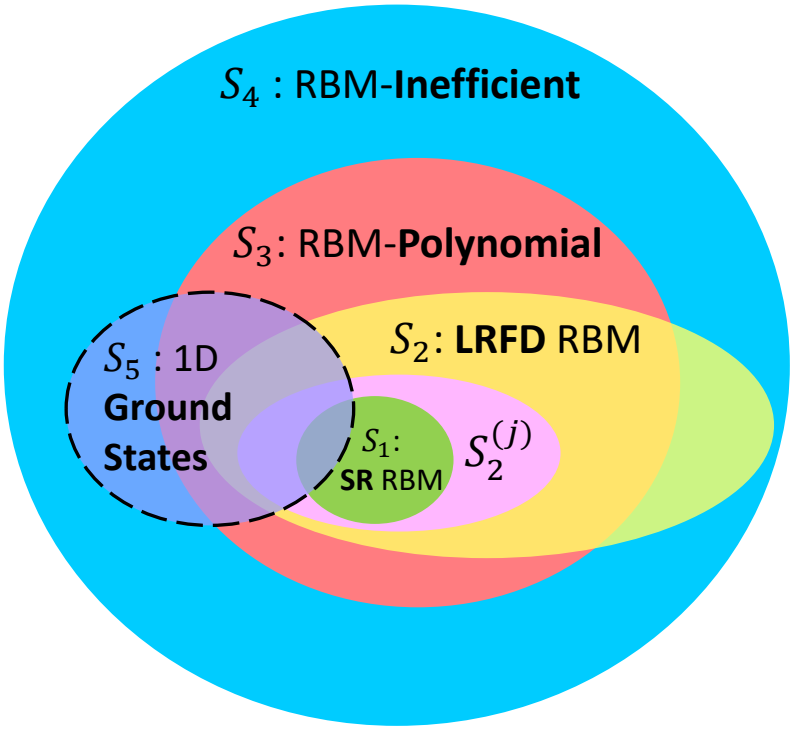


Figure 2.5: Relations between multiple typical state manifolds. S_1: short-range RBMs; S_2: LRFD RBMs; $S_2^{(j)}$ (for $1 \leq j \leq 6$, $j \in \mathbb{N}$): LRFD RBMs with distinct parameter conditions, specified in Table 2.1; S_3: RBMs with spatial complexities scaling at most polynomially in system sizes; S_4: RBMs with a faster-than-polynomial scaling of spatial complexities in system sizes, corresponding to inefficiency of representation; S_5: ground states of 1D quantum spin systems. The dashed boundary of S_5 means that its relations with other manifolds have not been fully determined.

manifold usually refers to a subspace of the whole Hilbert space spanned by a parameterized wave-function family [58], thus it is a set containing a specific scope of quantum states. The manifolds S_1, S_2, $S_2^{(j)}$ (for $1 \leq j \leq 6$, $j \in \mathbb{N}$), S_3 and S_4 are defined to be the space spanned by quantum states represented by RBMs satisfying corresponding conditions as given in Fig. 2.5, while S_5 is defined to be the manifold spanned by all ground states of 1D quantum many-body spin systems.

The definitions of these manifolds directly implies that $S_1 \subsetneq S_2^{(j)} \subsetneq S_2$ (for $1 \leq j \leq 6$). Our complexity analysis for LRFD RBMs (Sec. 2.2.3) gives the result that $S_2^{(j)} \subseteq (S_2 \cap S_3)$. Previous research shows that a set of problems where RBMs appear to be powerful are related to topological states, among which the 1D SPT cluster states belong to $S_5 \cap S_1$ [30]. The Laughlin wave functions, which have the structure of Jastrow wave functions and are associated with chiral topological order, can be exactly represented by RBMs in S_3 with a quadratic scaling of N_h in L but their approximations with RBMs of a long-range form and less complexity are often used [8]. S_4 contains all other sets mentioned in Fig. 2.5 as RBMs without restriction on the number of hidden nodes are universal approximators for discrete distribution [33]. Numerical results seem to support that a "large fraction" of S_5 is contained in its intersection with S_2. We argue that the concept of S_2 may benefit the understanding of which fraction of S_5 falls into its intersection with S_3, thus also promoting the understanding of the complexity of quantum many-body states.

It is remarkable that our paradigm for complexity analysis and our characterization of the nonlocal structures of RBMs for 1D quantum spin systems can be generalized to higher-dimensional systems, e.g., lattices. This is done by generalizing the description of single-level "orbitals" from $\mu(r)$ to $\mu(\vec{r})$ while keeping $\lambda(\tilde{k})$ as a level-decay factor. For deep

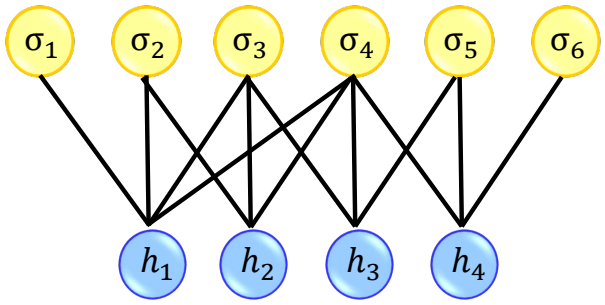


Figure 2.6: Network structure of a sparse RBM to be transformed into an MPS. This RBM serves as an example to show the possible failing of the original transformation algorithm and the effect of our improvement.

NN quantum states, we can still view each single hidden layer as a combination of multiple levels which capture correlations of different orders. We can calculate the truncation errors for each hidden layer associated with specific nodal functions and analyze the propagation of errors through layers.

2.5 Transformation from RBMs to MPSs

There are some works studying the relationship between RBMs and other concepts about state representations, such as string-bond states [8], correlator product states [55] and tensor network states [38, 56]. Especially, the transformation from RBMs to MPSs is analyzed [56]. But such types of transformations may lead to redundancy in parametrization when the RBMs are not quite short-range or sparse as they only use structural information of the network. An algorithm for finding an optimal mapping from RBMs to MPSs is given in Ref. [56], but we point out that this algorithm may fail and mainly works for short-range or very sparse RBMs.

The core idea of the optimal transformation algorithm is to find a minimal-size "separation set" when dichotomizing the visible nodes into two parts. In other words,

when the nodes in the "separation set" are excluded, the remaining two subsets of visible nodes are not connected. It implies that the wave function of the quantum state can be factorized into two independent parts conditioned on fixed values of spins in the "separation set". This "separation set" is very similar to the concept of a "Markov blanket" in the undirected graph, but is different from the latter concept in that it allows the nodes in the separated sets to be included.

We explore the reason for the failure of the optimal transformation algorithm and put forward an improvement of the algorithm. The algorithm will fail if there are common visible nodes in the two node sets to be separated ("X" and "Y" sets in Ref. [56]), because they should be included in the "separation set", but can not be done so as their corresponding indices have been contracted at the last step. Such a phenomenon will appear when the "separation set" at some step of the tensor construction contains too many visible nodes waiting to be contracted. So this algorithm will fail when the RBM is not so sparse. One example of the failing of this algorithm is given in Fig. 2.6 and Table 2.2. We point out that this algorithm can be improved by not randomly choosing the "separation set" at the construction of tensor $A[\sigma_j]$ when there are several choices, but rather choosing the one containing "σ_j" if possible. This principle can extend the length of steps in which there exists a separation set. The effect of adopting such a principle is given in Table 2.3. The improved algorithm can generate an MPS while the original one fails for the same RBM. In both tables, w_{jk} is an abbreviation of $W_{jk}\sigma_j h_k$.

step	X	Y	separation set	tensor
1	σ_1	$\sigma_2, ..., \sigma_6$	h_1	$A^{(1)}_{1,h_1}[\sigma_1] = \exp(a_1\sigma_1 + w_{11})$
2	h_1, σ_2	$\sigma_3, ..., \sigma_6$	σ_3, σ_4	Algorithm fails

Table 2.2: Failing of the algorithm of finding an optimal transformation from RBM to MPS. The RBM to be transformed has a network structure shown in Fig. 2.6. The original algorithm uses a random selection of "separation sets".

step	X	Y	separation set	tensor
1	σ_1	$\sigma_2, ..., \sigma_6$	σ_1	$A^{(1)}_{1,\sigma_1}[\sigma_1] = \exp(a_1\sigma_1)$
2	σ_1, σ_2	$\sigma_3, ..., \sigma_6$	σ_2, h_1	$A^{(2)}_{\sigma_1,\sigma_2 h_1}[\sigma_2] = \exp(a_2\sigma_2 + w_{11} + w_{21})$
3	σ_2, h_1, σ_3	$\sigma_4, ..., \sigma_6$	σ_3, σ_4	$A^{(3)}_{\sigma_2 h_1,\sigma_3\sigma_4}[\sigma_3] = \sum_{h_1,h_2} \exp(a_3\sigma_3 + w_{22} + w_{31} + w_{32} + w_{41} + w_{42} + b_1h_1 + b_2h_2)$
4	σ_3, σ_4	σ_5, σ_6	h_3, h_4	$A^{(4)}_{\sigma_3\sigma_4,h_3\sigma_4}[\sigma_4] = \exp(a_4\sigma_4 + w_{33} + w_{43} + w_{44})$
5	h_3, h_4, σ_5	σ_6	σ_6	$A^{(5)}_{h_3h_4,\sigma_6}[\sigma_5] = \sum_{h_3,h_4} \exp(a_5\sigma_5 + w_{53} + w_{54} + w_{64} + b_3h_3 + b_4h_4)$
6	σ_6	$\varnothing$	$\varnothing$	$A^{(6)}_{\sigma_6,1}[\sigma_6] = \exp(a_6\sigma_6)$

Table 2.3: Success of generating an MPS using our improved algorithm. The RBM to be transformed is the same as that in Table 2.2.

2.6 Conclusion

In this work, we define a subset of generic RBM quantum states—long-range-fast-decay (LRFD) RBM states. Using the language of levels, the nonlocal structure of LRFD RBMs is described with two functions: one of which, $\mu(r)$, captures the localization of the spatial distribution of the wave function for each single level and encodes information about spin correlations; the other, $\lambda(\tilde{k})$, is a level-decay factor capturing correlations of different orders and significantly influencing the complexity of the RBMs. We derive upper bounds on truncation errors, which allow us to analyze the scaling of the spatial complexity in system sizes and approximation errors for LRFD RBMs. We provide numerical results supporting that the ground states of a wide range of 1D quantum spin systems, including

some critical systems, may be approximated by LRFD RBMs with an at-most-polynomial complexity. Finally, we describe the relationships between state manifolds of different computational complexity and identify hierarchies of RBM-efficient approximation.

Generalizing the RBM wave-function ansatz to an infinitely-many-hidden-node regime and proposing the concept of LRFD RBMs does not imply the use of an infinitely-large neural network for state representations. These serve to define the completeness of a set of variational states and serve as a tool for complexity analysis based on the good extensibility and analyzability of LRFD-RBM forms. This concept may promote general understanding of the intrinsic complexity of quantum many-body states.

Chapter 3: Learning quasiparticle excitations in long-range interacting quantum systems using neural networks

3.1 Introduction

The interface between machine learning and quantum information processing is an emerging and rapidly advancing field [2, 3, 5, 6, 24, 72, 73]. Besides the research on quantum algorithms [11, 12, 13, 14] and physics-inspired ideas to enhance traditional machine learning [15, 16] and on the fundamental limitations of physical agents based on quantum information theory [74, 75], there has also been tremendous progress in directly applying machine learning techniques to quantum systems. Indeed, these techniques were applied to a wide range of topics including the identification of quantum phases and transitions [3, 17, 18, 19, 21], molecular modeling [22, 23], and quantum state tomography [24, 25]. Neural network states have been extensively studied recently as they provide a compact wave-function ansatz in the context of quantum many-body physics [5, 6, 7, 8, 9, 28, 30, 31]. These states have achieved success in simulating the low-lying eigenstates and short-time unitary dynamics of quantum many-body-localized systems [6, 34] and representing particular states such as code words of a stabilizer code [30, 35, 36] and chiral topological states [8, 37].

Numerous atomic, molecular, and optical systems exhibiting long-range interactions are emerging as versatile platforms for quantum computation and quantum simulation [39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52]. These long-range interactions include dipolar (decaying as $1/r^3$ with distance r) interactions between electric [39, 40] or magnetic dipoles [41, 42, 43], strong van der Waals ($1/r^6$) interactions between Rydberg atoms [39, 44, 45] or Rydberg polaritons [46], along with more general forms of interaction between trapped ions ($1/r^\alpha$, $0 \leq \alpha \leq 3$) [47, 48, 49, 50, 51, 52]. Following a quench, quantum information in long-range interacting systems can propagate faster than in short-range interacting ones [48, 49, 76, 77, 78]. The notion of quasiparticles as elementary excitations often provides an effective way for understanding non-equilibrium dynamics and quantum thermalization [79, 80, 81]. For instance, the dynamics of correlation spreading and entanglement growth are often constrained by an effective light cone and at low energy they can often be ascribed to the propagation of quasiparticles through the system [76, 77, 82, 83, 84, 85, 86, 87, 88, 89, 90]. This light cone is linear ($t \sim r$) for short-range interactions, but, for small α, can become sub-linear (e.g., $t \sim r^\beta$, $0 < \beta < 1$) or even instantaneous (for any r, $t = 0$ in the thermodynamic limit). In the case of linear light cones, the maximum speed of quantum information propagation can often be related to the maximum group velocity of quasiparticles, while the sub-linear and instantaneous light cones can often be associated with infinite group velocities of quasiparticles in the thermodynamic limit. In higher dimensions D and in the presence of strongly long-range interactions, it can be hard to numerically study the system with some widely-adopted methods, such as the density matrix renormalization group, due to the violation of entanglement area law [69, 91, 92] and intractability of contracting higher-dimensional projected-entangled pair states [93].

In this work, we use restricted Boltzmann machines (RBMs) to learn the momentum-resolved low-energy spectrum of long-range interacting systems, which are studied in current quantum simulation experiments [47, 48, 49, 50, 51, 52]. We introduce an energy-shift method to calculate the excited states. The core idea of the method is to shift the energies of the original Hamiltonian in a way that makes a target excited state the ground state of the new Hamiltonian. This method has an advantage that the target excited state is represented by a single RBM along the entire optimization path, rather than by multiple RBMs [34]. Such a single-RBM representation has a smaller parameter space and its variational parameters have a more natural physical interpretation, so that the geometrical properties, such as the quantum Fisher matrix, of the optimization path can be studied more easily and help understand the corresponding quantum phase [94, 95, 96]. Combined with a fixed-quantum-number method [34] generalized to long-range interactions, this method can resolve the full quasiparticle dispersion relation even in the presence of strongly long-range interactions and works independently of the ground-state degeneracy and of the size of the gap above the ground-state manifold.

Based on the low-energy spectrum, we identify the critical exponent α_c where the maximal quasiparticle group velocity transits from finite to divergent in the thermodynamic limit in $D = 1$. The results are quantitatively consistent with an analysis using field theory and linear spin-wave theory [68, 76, 85, 97]. Our use of RBMs to learn the spectrum of two-dimensional (2D) models is particularly noteworthy since other numerical methods, such as tensor network states, are often computationally intractable in 2D [93]. Our results can help understand the information propagation speed, entanglement growth, and the rate of thermalization in long-range interacting quantum systems. Our work can also be used

to provide benchmarks for large-scale 1D and 2D state-of-the-art quantum simulators.

3.2 Learning excited states with RBMs

We use the RBM as a variational wave-function ansatz to learn the eigenstates of quantum many-body spin systems. An RBM is a stochastic artificial neural network for probability modeling defined on a bipartite undirected graph [6]. Its first layer is the visible layer containing a spin configuration $\vec{\sigma}$ as input, where $\vec{\sigma} = (\sigma_1, \ldots, \sigma_L)$ has L degrees of freedom and $\sigma_j = \pm 1$ $(j = 1, \ldots, L)$ for spin-1/2 systems. The second layer is comprised of hidden nodes $\vec{h} = (h_1, \ldots, h_{N_h})$ ($h_k = \pm 1$ for $k = 1, \ldots, N_h$), which are introduced as auxiliary spins in the probability model. Given a specific spin configuration $\vec{\sigma}$, the RBM outputs a wave function amplitude:

$$\begin{aligned} \psi(\vec{\sigma}) &= \sum_{\vec{h}} e^{\sum_j a_j \sigma_j + \sum_k b_k h_k + \sum_{j,k} W_{jk} \sigma_j h_k} \\ &= \prod_{j=1}^{L} e^{a_j \sigma_j} \prod_{k=1}^{N_h} 2\cosh(b_k + \sum_j \sigma_j W_{jk}), \end{aligned} \tag{3.1}$$

where a_j and b_k are the visible and hidden biases, respectively, W_{jk} are weights corresponding to interlayer interactions, and all these parameters are complex numbers. All such amplitudes defined on the computational basis yield a quantum state vector $|\Psi\rangle = \sum_{\vec{\sigma}} \psi(\vec{\sigma})|\vec{\sigma}\rangle$. This RBM ansatz possesses good expressive power [7] and can often be optimized (trained) by several well-developed numerical methods, such as the stochastic reconfiguration method [6, 64, 65] and stochastic gradient descent [66].

While the RBM can be used to find ground states by reinforcement learning [6], its

applicability to low-lying excited states of various quantum many-body systems and the development of efficient optimization methods are less fully explored. Here we introduce an energy-shift method which shifts the energies of the original Hamiltonian by adding projection operators to make a target excited state the ground state of the new Hamiltonian. Let $|\Psi_j\rangle$ denote the j-th excited states of a given Hamiltonian $\hat{H}_0$ for a system containing L spins, where $j = 0, 1, \ldots, 2^L - 1$ and $j = 0$ corresponds to the ground state. To calculate $|\Psi_j\rangle$, the idea is to first calculate all the lower excited states $|\Psi_l\rangle$ $(0 \leq l \leq j-1)$ and then modify $\hat{H}_0$ by adding projection operators that lift the energies of all $|\Psi_l\rangle$:

$$\hat{H}_j = \hat{H}_0 + \sum_{l=0}^{j-1} E_l^{(\text{shift})} \frac{|\Psi_l\rangle\langle\Psi_l|}{\langle\Psi_l|\Psi_l\rangle}. \tag{3.2}$$

This treatment is analogous to the quadratic penalty method for the optimization problem with equality constraints [98] and leads to the convergence of the RBM to $|\Psi_j\rangle$ in the effective imaginary-time evolution [6], provided that $E_l^{(\text{shift})}$ is larger than the energy difference between $|\Psi_j\rangle$ and $|\Psi_l\rangle$. These projection operators can be efficiently implemented (Appendix B.1), thus at least low-lying excited states can be calculated successively in principle. This method is also applicable to resolving nearly degenerate ground states. But for systems with exact ground-state degeneracy such as spin glasses [99] and topologically ordered systems [100, 101], how to improve the accuracy under the sign problem [102] and whether RBMs can properly diagonalize the degenerate subspace deserves more investigations.

For translationally invariant systems, the symmetry information can be utilized to calculate the lowest-energy eigenstate in each momentum sector [34]. In this RBM variant, the original RBM wave-function ansatz (Eq. (4.1)) only applies to canonical spin configu-

rations, while the amplitudes of other configurations are constructed by building mappings onto canonical configurations so that the whole state satisfies translational symmetry constraints. Specifically, in $D = 1$, let $\hat{T}$ denote the unit-distance translation operators for spin configurations along the chain and define any configuration $\vec{\sigma}$ with all the configurations obtained by applying $\hat{T}$ to $\vec{\sigma}$ multiple times as one equivalence class. Then the set of canonical spin configurations $S_{\text{canonical}}$ is established by picking one element from each such equivalence class. Define the amplitude of $\vec{\sigma}$ to be

$$\tilde{\psi}(\vec{\sigma}) = e^{ikr_{\vec{\sigma}}}\psi(\vec{\sigma}_{\text{canonical}}), \tag{3.3}$$

where $k = \frac{2\pi m}{L}$ (for $m = 0, 1, \ldots, L-1$) is momentum, ψ has the original RBM form (Eq. (4.1)), $\vec{\sigma}_{\text{canonical}}$ is the canonical configuration in the same equivalence class as $\vec{\sigma}$ and they are associated by $\vec{\sigma} = \hat{T}^{r_{\vec{\sigma}}}\vec{\sigma}_{\text{canonical}}$. The lowest-energy state with the above wavefunction ansatz is automatically the eigenstate of $\hat{T}$ with eigenvalue e^{ik}, thus just the lowest excited state in the sector of momentum k.

A quasiparticle dispersion relation curve can be obtained by using both our energy-shift method and the fixed-quantum-number method regardless of whether the model is gapped or gapless. For gapless models, we use the fixed-quantum-number method to calculate the lowest excited states located in momentum sectors different from the ground state. For gapped models, we do additional calculations for the lowest excited state in the same momentum sector as the ground state but separated by a gap that will not vanish in the thermodynamic limit using the energy-shift method.

We first use these numerical methods to study two 1D power-law interacting quantum

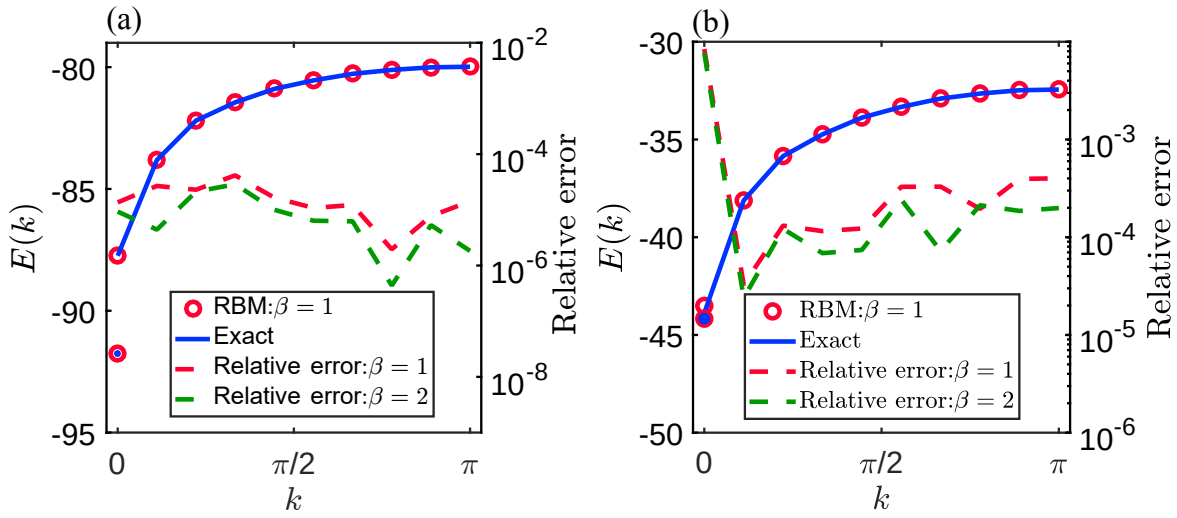


Figure 3.1: Low-energy spectra (red circles) obtained by RBMs with hidden-unit density $\beta = 1$ compared to the results obtained by exact diagonalization (solid blue line), and the relative errors on energy values using RBMs with $\beta = 1$ and 2 (dashed lines). (a) Gapped long-range TFIM with $\alpha = 1.2$, $B = 5$ and $L = 18$. (b) Gapless long-range XXZ model with $\alpha = 1.2$, $J_z = -0.5$ and $L = 18$.

spin systems, which have been engineered in quantum simulation experiments [47, 48, 49]. Their Hamiltonians are described by the following long-range interacting transverse-field Ising model (TFIM) [Eq. (4.4)] and XXZ model [Eq. (3.5)]:

$$\hat{H} = -\sum_{1\le j<k\le L} \frac{\hat{\sigma}_j^z \hat{\sigma}_k^z}{r_{jk}^\alpha} - B\sum_{j=1}^{L} \hat{\sigma}_j^x, \tag{3.4}$$

$$\hat{H} = \sum_{1\le j<k\le L} \frac{1}{r_{jk}^\alpha}(-\hat{\sigma}_j^x \hat{\sigma}_k^x - \hat{\sigma}_j^y \hat{\sigma}_k^y + J_z \hat{\sigma}_j^z \hat{\sigma}_k^z), \tag{3.5}$$

where $\hat{\sigma}_j^\mu$ denotes the spin-1/2 Pauli operator at site j of the chain ($\mu = x, y, z$), L is the system size, $r_{jk} = \min(|j-k|, L-|j-k|)$ as we use periodic boundary conditions. α is a positive constant for describing the power-law interactions and we only focus on $\alpha > D$ in this work to ensure a well-behaved thermodynamic limit without the necessity of Hamiltonian normalization. B is a global transverse field and J_z is the coupling constant for

the z spin components. The ground state of this long-range TFIM with fixed ferromagnetic interactions undergoes a phase transition from a gapped ordered ferromagnetic phase with broken $\mathbb{Z}_2$ symmetry to a gapped disordered paramagnetic phase. This phase transition is similar to that for a short-range TFIM but at a field strength B which increases with decreasing α [67, 103]. The ground state of the long-range XXZ model exhibits four phases. Besides gapped ferromagnetic and antiferromagnetic Ising phases for large $|J_z|$, there exists a gapless disordered XY phase known as Luttinger liquid for small J_z and large α, while a gapless continuous symmetry breaking phase appears for sufficiently small α which breaks the U(1) rotational symmetry around the z axis [68, 85].

Figs. 3.1(a) and (b) show the low-energy spectra of the two models computed by the RBM learning. We benchmark our method by comparing to exact-diagonalization results. The accuracy quantified by the relative error on energy values generally improves with increasing hidden-unit density $\beta = N_h/L$ while for $\beta = 1$ the learning has already achieved high accuracy with relative errors smaller than 10^{-4} for TFIM and 4×10^{-4} for XXZ model, respectively, except for the first-excited state of XXZ model which has a larger relative error. The accuracy of the latter gapless model is slightly lower perhaps because the system is in the continuous symmetry breaking phase with long-range order so that more hidden nodes are required to capture the nonlocal correlations between spins [6]. It is also remarkable that the generalization of the so-called local energy to include nonlocal terms induced by long-range interactions in the stochastic reconfiguration method seems not to complicate the learning algorithm explicitly, but influence the quantum Fisher matrix and thus the learning landscape [95] in an implicit way (Appendix B.1).

3.3 Group velocity transition

In long-range interacting systems, quench dynamics can present different regimes based on whether locality persists or is broken [48, 49, 76, 77]. Quasiparticles as elementary excitations often provide an effective way for understanding such low-energy dynamics. The quasiparticle propagating with the maximum group velocity can often be thought of as forming the front of the spreading correlations. This quasiparticle then also determines the speed of quantum information propagation and entanglement growth in these systems [76, 77, 83, 84, 85, 86]. An infinite speed of information propagation may arise for α smaller than some critical value in long-range interacting systems, which can often be related to the divergence of the maximum group velocity of quasiparticles in the thermodynamic limit.

Here we use our RBM learning method to compute the dispersion relation of quasiparticles in order to identify the value of the critical exponent α_c. We apply RBM learning to the models in Eqs. (4.4) and (3.5) for a range of α and for system sizes up to $L = 70$. For the long-range TFIM, a cusp structure (a nonanalyticity) appears at momentum $k = 0$ for small α [Fig. 3.2(a)], which implies the divergence of the group velocities for quasiparticles with $k \to 0$. For large α, the dispersion curve has a finite slope everywhere, which implies bounded maximum group velocities of quasiparticles in the thermodynamic limit. For large α, the group velocity doesn't attain its maximum, but rather approaches zero, as $k \to 0$. For the long-range XXZ model, the dispersion relations have similar features as in the long-range TFIM, but at large α, they are almost straight lines with finite slopes near $k = 0$, corresponding to the usual linear dispersion of Goldstone modes [85] [see Fig. 3.2(b)].

To locate the critical α, we then compute the group velocities at the momentum $k = 0$

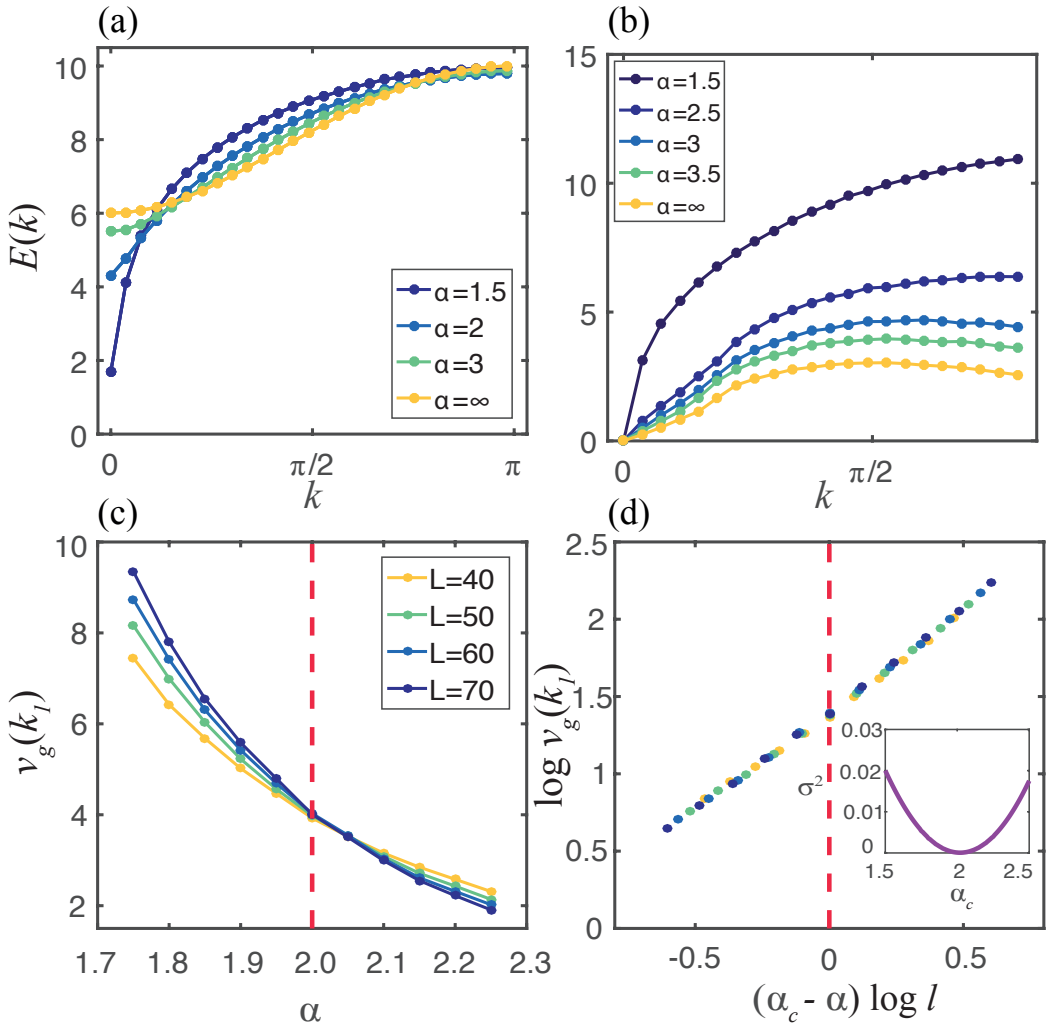


Figure 3.2: (a,b) Dispersion relations shifted vertically so that the ground-state energy is zero. The results are obtained by RBMs with $\beta = 1$. (a) TFIM with $\alpha = 1.5$, 2, 3 and ∞, while $B = 4$ and $L = 53$. (b) XXZ model with $\alpha = 1.5$, 2.5, 3, 3.5 and ∞, while $J_z = -0.5$ and $L = 53$. (c) Group velocity $v_g(k_1)$ as a function of α for different system sizes L for TFIM with $B = 4$. (d) Data collapse for (c). The insets in the lower right corner show the variance in data collapse processing for a range of α_c.

where a cusp may emerge [Figs. 3.2(c)]. $v_g(k=0)$ is calculated using a finite-difference form $v_g(0) = (E(\frac{2\pi}{L}) - E(0))/(\frac{2\pi}{L})$, where $E(\frac{2\pi}{L})$ is the lowest energy in the $k = \frac{2\pi}{L}$ momentum sector, while $E(0)$ is the energy of the first excited state (ground state) located in the $k = 0$ sector for gapped (gapless) models. For the long-range TFIM [Figs. 3.2(c)], the curves of group velocity $v_g(k_1)$ for different system sizes cross at a critical $\alpha = \alpha_c$ near 2. For $\alpha < \alpha_c$, $v_g(k_1)$ increases with growing system size, while the monotonic trend is reversed for $\alpha > \alpha_c$. In the thermodynamic limit, we expect the curve to develop a non-analyticity at $\alpha = \alpha_c$ with divergent (vanishing) $v_g(k_1)$ below (above) α_c.

In order to get the precise value of α_c in an algorithmic way, we perform a finite-size data collapse [104] [Figs. 3.2(d)] by plotting $(\alpha_c - \alpha)\log l$ and $\log v_g(k_1)$ data pairs, where $l = \frac{L}{2\pi} = \frac{1}{k_1}$. The data for the long-range TFIM [Fig. 3.2(e)] collapses onto a scaling function, which approaches the positive (negative) infinity when extending along the positive (negative) direction, consistent with the expected behavior of $v_g(k_1)$ across $\alpha = \alpha_c$ in the thermodynamic limit. The results show good data collapse when fitting with $\alpha_c = 2$ for the TFIM model, and the scaling functions are nearly linear. This justifies the validity of our finite-size analysis and implies that the dominant term in $v_g(k_1)$ has a power-law dependence on $l^{\alpha_c - \alpha}$ for the collapsing range of l and α. We further define the variance σ^2 to quantitatively measure the deviation in data collapse for different choices of α_c [see insets of Figs. 3.2(d)]. This analysis shows that $\alpha_c = 2.02$ gives the best collapse for the model.

Our numerical results are consistent with a field-theoretic analysis, at $\alpha > 1$, of

dispersion relations in the long-wavelength limit [85, 97]:

$$\omega(k) = \sqrt{\Delta^2 + ak^{\alpha-1} + bk^2}, \text{ as } k \to 0, \tag{3.6}$$

where $\omega(k)$ is the excitation energy, Δ is the gap, while a and b are model-dependent parameters. This formula yields $\alpha_c = 2$ [97] and 3 [85] for the gapped ($\Delta \neq 0$) TFIM and the gapless ($\Delta = 0$) XXZ model, respectively, and gives the finite-size behavior of $v_g(k_1)$ across $\alpha = \alpha_c$ that agrees well with the above results. We note that, in the above parameter regimes, the long-range TFIM stays in the gapped paramagnetic phase [67], while the XXZ model stays gapless but undergoes a phase transition from the XY phase to continuous symmetry breaking phase [68]. Linear spin-wave theory, which is applicable to systems deep in these phases, also gives the same theoretical predication of α_c [76, 85], in agreement with our results for both models.

3.4 Correlations in excited states

With well-trained RBMs representing the excited states, we are able to efficiently investigate, via Monte-Carlo sampling, a range of properties of these states, such as the correlation decay. In the paramagnetic phase, the longitudinal correlations $\langle \hat{\sigma}_1^z \hat{\sigma}_{1+r}^z \rangle$ in the ground state of the TFIM have a power-law decay down to zero [88, 105]. However, we see that longitudinal correlations in the first excited state with $k = 0$ have an almost power-law decay at short range, but converge to a nonzero constant proportional to $\frac{1}{L}$ [Fig. 3.3(a)], whereas other low-lying excited states with $k \neq 0$ behave similarly to the ground state. The vanishing long-distance correlations in the thermodynamic limit imply that there is

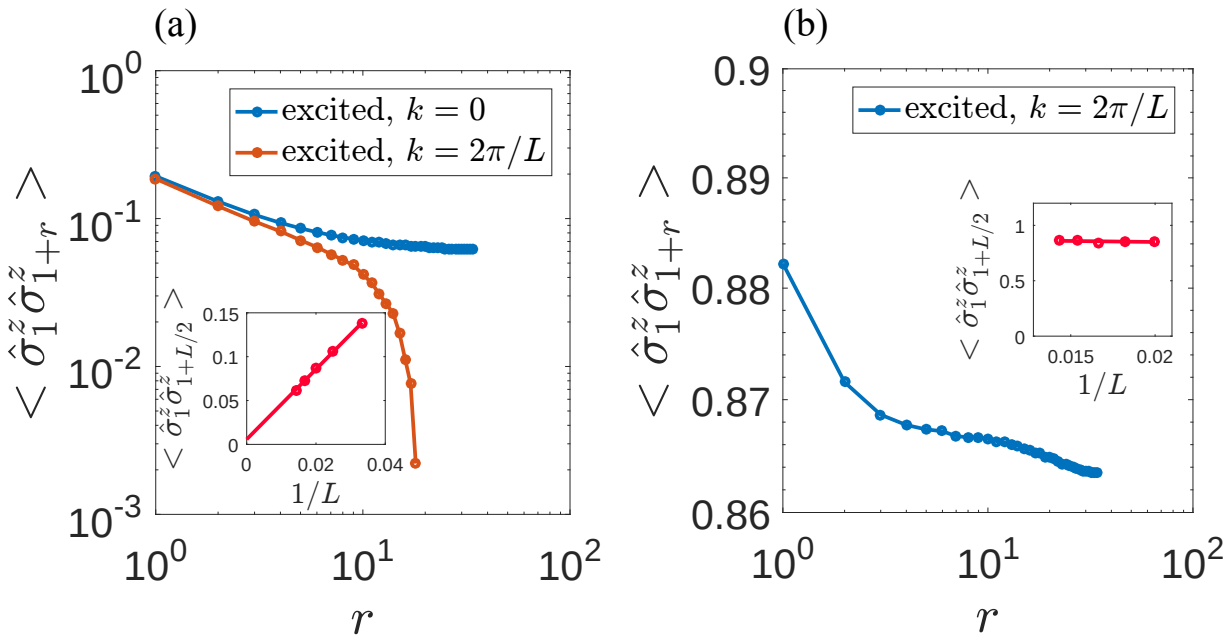


Figure 3.3: Longitudinal correlations in low-lying excited states of the long-range TFIM in different phases. (a) The lowest excited states in sectors $k = 0$ and $k = 2\pi/L$ in the paramagnetic phase with $\alpha = 1.8$, $B = 4$ and $L = 70$. Inset: the long-distance limits of longitudinal correlations in the lowest excited state in the sector $k = 0$ for $30 \leq L \leq 70$. (b) The lowest excited state in the sector $k = 2\pi/L$ in the ferromagnetic phase with $\alpha = 1.8$, $B = 0.9$ and $L = 70$. Inset: the long-distance limits of longitudinal correlations in the lowest excited state in the sector $k = 2\pi/L$ for $50 \leq L \leq 70$.

also no long-range order in the low-lying excited states when the ground state of TFIM is in the paramagnetic phase. By contrast, when the ground state is in the ferromagnetic phase, the longitudinal correlations in the low-lying excited states converge to constants which do not decay to zero as system size L grows [Fig. 3.3(b)] indicating the existence of long-range order. This is consistent with the intuition that the presence of single quasiparticles does not make the low-lying excited states have an ordered or disordered nature different from the ground states. Therefore, the long-distance limit of correlation functions $\langle\hat{\sigma}_1^z\hat{\sigma}_{1+L/2}^z\rangle$ in the low-lying excited states of the TFIM presents distinct behaviors across the α-dependent phase transition point $B(\alpha)$. This feature of excited states can be utilized to compute the phase boundary in the B-versus-α plane by extracting the behavior-transition points, which will be justified in the future study.

3.5 2D generalization

While the high-dimensional generalizations of some numerical methods, such as tensor network states, are often intractable [93], we can generalize the RBM method straightforwardly to learn the many-body excitations of 2D models. Specifically, we need to construct a wave-function ansatz satisfying translational symmetry constraints along two independent directions associated with lattice geometries. Here we study 2D long-range TFIMs on square and triangular lattices, the latter of which have been implemented in some quantum simulation experiments [45, 50, 51]. While the ground-state algorithm and the energy-shift method for excited states can be applied to any dimension directly, the crucial points when generalizing the fixed-quantum-number method are the extended definition of equivalence

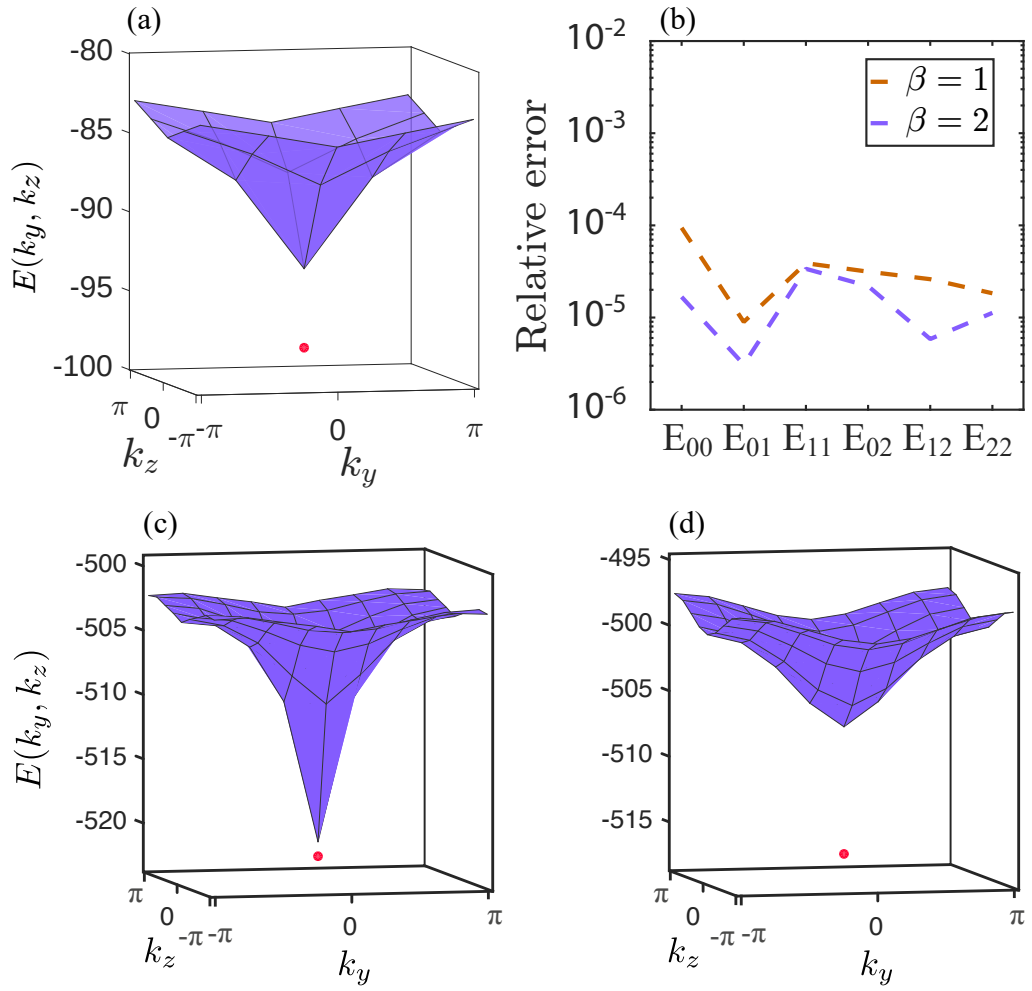


Figure 3.4: 2D long-range TFIM on a square lattice. (a) Low-lying energy spectrum by generalized RBMs with hidden-unit density $\beta = 2$. $\alpha = 3.5$, $B = 6$, $L = 4 \times 4 = 16$. (b) Relative error in energy calculations in (a) compared to exact-diagonalization solutions. $E_{j_1 j_2}$ denotes $E(k_y = \frac{2\pi}{\sqrt{L}} j_1, k_z = \frac{2\pi}{\sqrt{L}} j_2)$. (c,d) Energy spectra by generalized RBM with $\beta = 1$. $\alpha = 2.5$ (for (c)) and $\alpha = 4.5$ (for (d)), $B = 8$, $L = 8 \times 8 = 64$. In (a), (c) and (d), the red points correspond to the ground states.

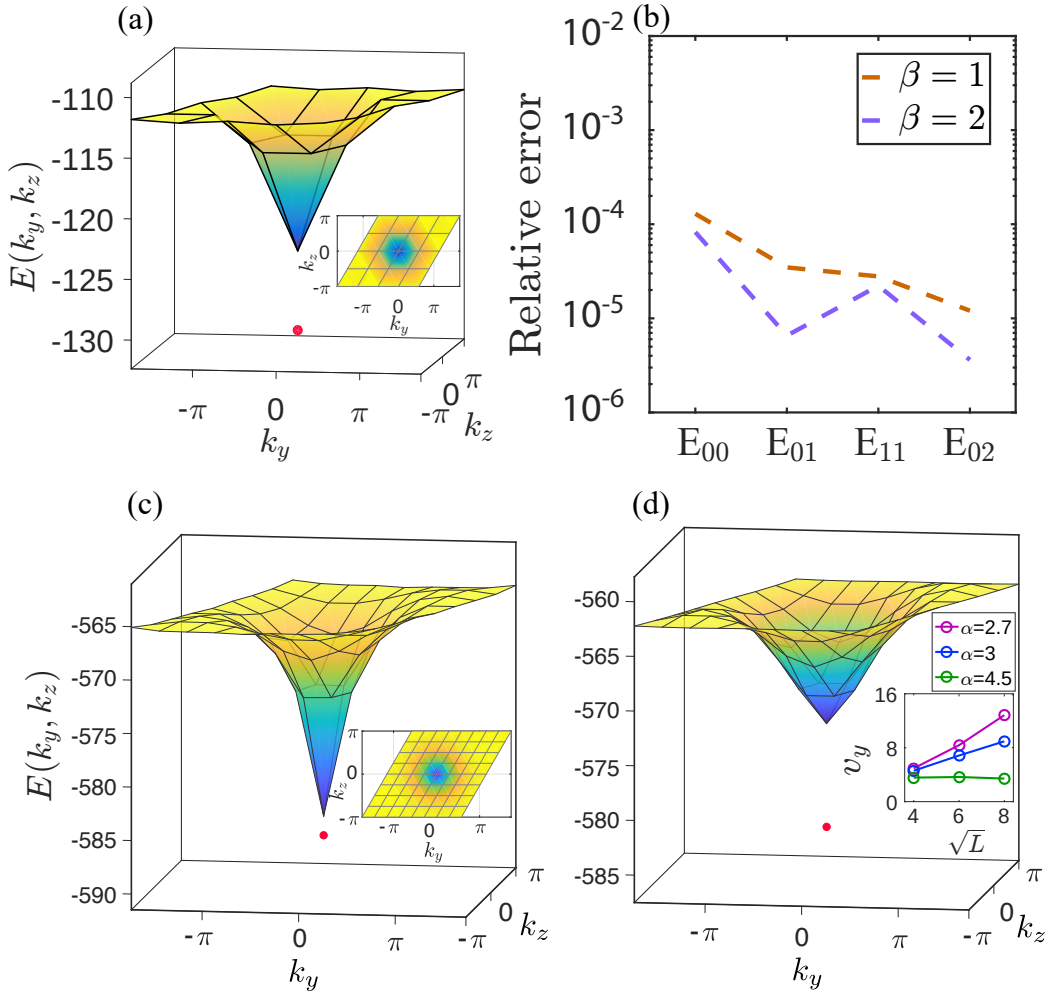


Figure 3.5: 2D long-range TFIM on a triangular lattice. (a) Low-lying energy spectrum by generalized RBMs with $\beta = 2$. $\alpha = 3$, $B = 8$, $L = 4 \times 4 = 16$. (b) Relative error in energy calculations in (a) compared to exact-diagonalization solutions. $E_{j_1 j_2}$ denotes $E(k_1 = \frac{2\pi}{\sqrt{L}} j_1, k_2 = \frac{2\pi}{\sqrt{L}} j_2)$. (c,d) Energy spectra by generalized RBMs with $\beta = 1$. $\alpha = 2.7$ (for (c)) and $\alpha = 4.5$ (for (d)), $B = 9$, $L = 8 \times 8 = 64$. In (a,c,d), the red points correspond to the ground states. In (a,c), the insets are the heat map of the corresponding energy spectra and exhibit the expected rotational symmetry. In (d), the inset is the scaling of the y-component v_y of group velocity at $\vec{k} = \vec{0}$ with system sizes ($L = 16$, 36, and 64) for different α by RBM learning.

classes and the proper selection of canonical spin configurations which should lead to good convergence and numerical stability [34].

At first, we consider a 2D long-range TFIM on a rectangular lattice with a Hamiltonian

$$\hat{H} = -\frac{1}{2}\sum_{\vec{j}\neq\vec{j'}} \frac{\hat{\sigma}^z_{\vec{j}}\hat{\sigma}^z_{\vec{j'}}}{r^{\alpha}_{\vec{j},\vec{j'}}} - B\sum_{\vec{j}} \hat{\sigma}^x_{\vec{j}}. \tag{3.7}$$

The lattice is in the $y-z$ plane with side length L_y (L_z) in the y (z) direction. $\vec{j} = (j_y, j_z)$ and $\vec{j'} = (j'_y, j'_z)$ denote the positions of spins where $j_y, j'_y \in \{1, 2, \ldots, L_y\}$, $j_z, j'_z \in \{1, 2, \ldots, L_z\}$. $r_{\vec{j},\vec{j'}} = \left[(\min\{|j_y - j'_y|, L_y - |j_y - j'_y|\})^2 + (\min\{|j_z - j'_z|, L_z - |j_z - j'_z|\})^2\right]^{\frac{1}{2}}$ denotes the distance between the two sites $\vec{j}$ and $\vec{j'}$ as the periodic boundary conditions are used. Such a lattice has two primitive vectors $\vec{a}_1 = (1, 0)$ (positive y direction) and $\vec{a}_2 = (0, 1)$ (positive z direction). Let $\hat{T}_y$ and $\hat{T}_z$ denote the unit-distance translation operators for spin configurations along $\vec{a}_1$ and $\vec{a}_2$ directions, respectively, and then periodic boundary conditions imply $\hat{T}_y^{L_y} = 1$ and $\hat{T}_z^{L_z} = 1$. Then energy eigenstates are eigenstates of $\hat{T}_y$ and $\hat{T}_z$ with eigenvalues e^{ik_y} and e^{ik_z}, where $k_y = \frac{2\pi m_1}{L_y}$ and $k_z = \frac{2\pi m_2}{L_z}$ (for $m_1 \in \{0, 1, \ldots, L_y - 1\}, m_2 \in \{0, 1, \ldots, L_z - 1\}$) specify the momentum. Then each equivalence class can be defined by $\{\hat{T}_y^{j_1}\hat{T}_z^{j_2}\Sigma : j_1 \in \{0, 1, \ldots, L_y - 1\}, j_2 \in \{0, 1, \ldots, L_z - 1\}\}$, where $\Sigma = \{\Sigma_{j_y,j_z} \in \{\pm 1\} : j_y \in \{1, 2, \ldots, L_y\}, j_z \in \{1, 2, \ldots, L_z\}\}$ denotes a specific 2D spin configuration and is one element in this class.

To facilitate the selection of canonical spin configurations, we map each 2D configuration to a unique binary number (for spin-$\frac{1}{2}$ systems) by rearranging it into a 1D bit array of length $L = L_y L_z$ by a fixed rule. In this work, we choose to rearrange Σ by aligning each row of spins sequentially in a single row to form a 1D array $\vec{\widetilde{\sigma}} = \{\widetilde{\sigma}_m =$

$\Sigma_{j_y,j_z} : m = (j_y - 1)L_z + j_z, 1 \leq m \leq L\}$. Finally, for each equivalence class, we select the lexicographically smallest one as the canonical configuration $\vec{\tilde{\sigma}}_{\text{canonical}}$ [106] and transform it reversely to find its corresponding 2D configuration $\Sigma_{\text{canonical}}$. Then the wave-function ansatz satisfying translational symmetry constraints, similar to 1D cases in Ref. [34], can be defined as

$$\phi(\Sigma) = \tilde{\psi}(\vec{\tilde{\sigma}}(\Sigma)) = e^{i\vec{k}\cdot\vec{r}(\Sigma)}\psi(\vec{\tilde{\sigma}}_{\text{canonical}}), \tag{3.8}$$

where $\vec{k} = (k_y, k_z)$ is momentum, $\vec{r}(\Sigma) = (r_y, r_z)$, ψ has the original RBM form (Eq. (4.1)), $\vec{\tilde{\sigma}}_{\text{canonical}}$ and $\Sigma_{\text{canonical}}$ are the 1D and 2D representations of the canonical configuration in the same equivalence class as Σ, respectively, and there is an association $\Sigma = \hat{T}_y^{r_y}\hat{T}_z^{r_z}\Sigma_{\text{canonical}}$.

Figs. 3.4(a) and (b) show the energy spectrum of this model learned by generalized RBMs and the relative error in energy calculations compared to exact-diagonalization solutions for small-size systems, respectively. The numerical results are of high accuracy with relative errors less than 10^{-4} even for hidden-unit density $\beta = 1$. It can be seen from Figs. 3.4(c) and (d) that the dispersion of such a model with $B = 8$ and of a large size clearly has a cusp for $\alpha = 2.5$, while the cusp almost flattens out for $\alpha = 4.5$.

Our generalized RBM method is also applicable to a spin model on a triangular lattice. Consider a triangular lattice with two primitive vectors $\vec{a}_1 = (\frac{\sqrt{3}}{2}, -\frac{1}{2})$ and $\vec{a}_2 = (0, 1)$ and the overall shape of the lattice is a parallelogram with side length L_1 in $\vec{a}_1$ direction and L_2 in $\vec{a}_2$ direction. The site position can still be denoted by (j_1, j_2) which corresponds to the crossing of the j_1th line along $\vec{a}_1$ and the j_2th line along $\vec{a}_2$. The two generators for equivalence classes can be defined by $\hat{T}_1$ and $\hat{T}_2$ which denote the unit-distance translation

operators for spin configurations along these new $\vec{a}_1$ and $\vec{a}_2$ directions, respectively. Then energy eigenstates are eigenstates of $\hat{T}_1$ and $\hat{T}_2$ with eigenvalues e^{ik_1} and e^{ik_2}, where $k_1 = \frac{2\pi m_1}{L_1}, k_2 = \frac{2\pi m_2}{L_2}, m_1 \in \{0, 1, \ldots, L_1 - 1\}, m_2 \in \{0, 1, \ldots, L_2 - 1\}$. Then the procedures of RBM generalizations are similar to that in the rectangular-lattice case, except that an additional transformation of k_1 and k_2 associated with two non-orthogonal directions to momentum k_y and k_z is needed. This TFIM also has a Hamiltonian of the form of Eq. (3.7), except that the appropriate triangular-lattice distance function $r_{\vec{j},\vec{j}'}$ is used.

The numerical results for the spectrum are also of high accuracy, as benchmarked by exact-diagonalization results for small system sizes (Figs. 4.5(a) and (b)). It can be seen from Figs. 4.5(c) and (d) that the dispersion of the 2D long-range TFIM on a triangular lattice with $B = 9$ has a sharp cusp for $\alpha = 2.7$, while the cusp becomes obtuse for $\alpha = 4.5$. From $\alpha = 2.7$ to $\alpha = 4.5$, there is a change of the scaling behavior of group velocity at $\vec{k} = \vec{0}$ with system sizes (inset of Fig. 4.5(d)), which justifies the presence of cusp structures for $\alpha = 2.7$ and the absence of cusp for $\alpha = 4.5$. The identification of the precise value of α_c for 2D models perhaps requires more computational resources and new protocols. Furthermore, we can compare our RBM learning method with some theoretical tools for investigating quantum magnets in 2D, such as the high-order series expansion method [107]. Our RBM method has advantages that it is not limited to the perturbative regime and may work well for all phases without assumptions about gap-closing conditions.

3.6 Summary of quadratic Hamiltonians with long-range hoppings and pairings

Fermionic quadratic Hamiltonians with long-range hoppings and pairings are of the form

$$\hat{H} = \sum_{r_i=1}^{L} J_0 b_{r_i}^\dagger b_{r_i} + \sum_{r_i=1}^{L} \sum_{\delta=1}^{\frac{L-1}{2}} \frac{1}{\delta^\alpha} [J_1 (b_{r_i}^\dagger b_{r_i+\delta} + b_{r_i+\delta}^\dagger b_{r_i}) + \Delta (b_{r_i} b_{r_i+\delta} + b_{r_i+\delta}^\dagger b_{r_i}^\dagger)], \tag{3.9}$$

where the subscript r_i denotes the lattice site position, $r_i + \delta$ for an operator denotes the position $\mod (r_i+\delta, L)$ considering periodic boundary conditions, J_0 denotes the same-site coupling strength, J_1 denotes strength for long-range couplings, Δ denotes the strength for long-range pairings, $J_0, J_1, \Delta \in \mathbb{R}$, J_1 and Δ are not zero simultaneously, and L is assumed to be odd without influencing the asymptotic analysis.

By rewriting the Hamiltonian with reation and annihilation operators associated with momentum k, we can obtain

$$\hat{H} = \sum_k [b_k^\dagger b_k (J_0 + J_1 \Gamma_k^{(\alpha)}) + b_{-k} b_k i\Delta f_k^{(\alpha)} - b_k^\dagger b_{-k}^\dagger i\Delta f_k^{(\alpha)}], \tag{3.10}$$

where

$$\Gamma_k^{(\alpha)} := 2 \sum_{\delta=1}^{\frac{L-1}{2}} \frac{\cos(k\delta)}{\delta^\alpha}, \qquad f_k^{(\alpha)} := \sum_{\delta=1}^{\frac{L-1}{2}} \frac{\sin(k\delta)}{\delta^\alpha}. \tag{3.11}$$

Then the dispersion relation (excitation energy as a function of momentum for different

momentum modes) is:

$$E_k = \sqrt{(J_0 + J_1\Gamma_k^{(\alpha)})^2 + 4\Delta^2(f_k^{(\alpha)})^2}. \tag{3.12}$$

Let α_c be the critical exponent where the group velocity $v_g(k) = \frac{dE_k}{dk}$ at $k = 0$ transits from finite to divergent in the thermodynamic limit.

The polylogarithm function $\mathrm{Li}_s(z)$ is defined as

$$\mathrm{Li}_s(z) = \sum_{k=1}^{\infty} \frac{z^k}{k^s}. \tag{3.13}$$

This definition is valid for arbitrary complex order s and for all complex arguments z with $|z| < 1$; it can be extended to $|z| \geq 1$ by the process of analytic continuation. Let $z = e^{\mu}$ and expand the function about $\mu = 0$. We have

$$\mathrm{Li}_s(e^{\mu}) = \Gamma(1-s)(-\mu)^{s-1} + \sum_{k=0}^{\infty} \frac{\zeta(s-k)}{k!}\mu^k. \tag{3.14}$$

This result holds for $|\mu| < 2\pi$ and, using the analytic continuation provided by the zeta functions, for all $s \neq 1, 2, 3, \ldots$. If the order is a positive integer, $s = n$, both the term with $k = n - 1$ and the gamma function become infinite, although their sum does not. But we can still get the expansion of the function using harmonic numbers. The above expansion form is enough for us to deduce α_c because we just need to analyze the asymptotic behaviors of the function as s approaches a positive integer.

Let $\mu = ik$, where $k \in \mathbb{R}$. Then, as $k \to 0$, for $\alpha \notin \mathbb{N}_+$,

$$\begin{aligned}
\mathrm{Re}(\mathrm{Li}_\alpha(e^{ik})) &\sim \zeta(\alpha) + \mathrm{Re}[\Gamma(1-\alpha)(-ik)^{\alpha-1}] + \frac{1}{2}\zeta(\alpha-2)(ik)^2 && (3.15)\\
&= \zeta(\alpha) + \Gamma(1-\alpha)\cos(\frac{3\pi}{2}(\alpha-1))k^{\alpha-1} - \frac{1}{2}\zeta(\alpha-2)k^2, && (3.16)\\
\mathrm{Im}(\mathrm{Li}_\alpha(e^{ik})) &\sim \Gamma(1-\alpha)\sin(\frac{3\pi}{2}(\alpha-1))k^{\alpha-1} + \zeta(\alpha-1)k. && (3.17)
\end{aligned}$$

Therefore,

$$\begin{aligned}
\Gamma_k^{(\alpha)} &= 2\,\mathrm{Re}(\mathrm{Li}_\alpha(e^{ik})) && (3.18)\\
&\sim a_1 + b_1 k^{\alpha-1} + c_1 k^2, && (3.19)\\
f_k^{(\alpha)} &= \mathrm{Im}(\mathrm{Li}_\alpha(e^{ik})) && (3.20)\\
&\sim a_2 k^{\alpha-1} + b_2 k, && (3.21)\\
E_k &\sim \sqrt{[J_0 + J_1(a_1 + b_1 k^{\alpha-1} + c_1 k^2)]^2 + 4\Delta^2(a_2 k^{\alpha-1} + b_2 k)^2} && (3.22)\\
&= \big\{(J_0 + J_1 a_1)^2 + 2(J_0 + J_1 a_1)J_1 b_1 k^{\alpha-1} + [J_1^2 b_1^2 + 4\Delta^2 a_2^2]k^{2\alpha-2} + 8\Delta^2 a_2 b_2 k^\alpha && (3.23)\\
&\quad + [2(J_0 + J_1 a_1)J_1 c_1 + 4\Delta^2 b_2^2]k^2 + 2J_1^2 b_1 c_1 k^{\alpha+1} + J_1^2 c_1^2 k^4\big\}^{1/2} && (3.24)
\end{aligned}$$

Therefore, the critical exponent α_c for generic fermionic quadratic Hamiltonians with long-range hoppings and pairings can be derived (Table 3.1).

Similarly, bosonic quadratic Hamiltonians with long-range hoppings and pairings are

Table 3.1: The critical exponent α_c for generic fermionic quadratic Hamiltonians with long-range hoppings and pairings. d_j with $1 \leq j \leq 7$ are constants dependent on J_0, J_1, Δ and α.

Gap/Gapless	Parameter conditions	E_k (as $k \to 0$)	α_c
gapped	$J_1 \neq 0$	$d_1 + d_2 k^{\alpha-1}$ $(1 < \alpha < 3)$ $d_1 + d_3 k^2$ $(\alpha > 3)$	2
gapped	$J_1 = 0$, $\Delta \neq 0$	$d_1 + d_4 k^{2\alpha-2}$ $(1 < \alpha < 2)$ $d_1 + d_3 k^2$ $(\alpha > 2)$	$\frac{3}{2}$
gapless	$\Delta \neq 0$	$d_5 k^{\alpha-1}$ $(1 < \alpha < 2)$ $d_6 k$ $(\alpha > 2)$	2
gapless	$\Delta = 0$, $J_1 \neq 0$	$d_5 k^{\alpha-1}$ $(1 < \alpha < 3)$ $d_7 k^2$ $(\alpha > 3)$	2

Table 3.2: The critical exponent α_c for generic bosonic quadratic Hamiltonians with long-range hoppings and pairings. d'_j with $1 \leq j \leq 9$ are constants dependent on J_0, J_1, Δ and α.

Gap/Gapless	Parameter conditions	E_k (as $k \to 0$)	α_c
gapped	$J_0 J_1 + (J_1^2 - \Delta^2) a_1 \neq 0$	$d'_1 + d'_2 k^{\alpha-1}$ $(1 < \alpha < 3)$ $d'_1 + d'_3 k^2$ $(\alpha > 3)$	2
gapped	$J_0 J_1 + (J_1^2 - \Delta^2) a_1 = 0$	$d'_1 + d'_4 k^{2\alpha-2}$ $(1 < \alpha < 3)$ $d'_1 + d'_5 k^4$ $(\alpha > 3)$	$\frac{3}{2}$
gapless	$\Delta \neq 0$	$d'_6 k^{(\alpha-1)/2}$ $(1 < \alpha < 3)$ $d'_7 k$ $(\alpha > 3)$	3
gapless	$\Delta = 0$, $J_1 \neq 0$	$d'_8 k^{\alpha-1}$ $(1 < \alpha < 3)$ $d'_9 k^2$ $(\alpha > 3)$	2

of the form

$$\hat{H} = \sum_{r_i=1}^{L} J_0 b_{r_i}^\dagger b_{r_i} + \sum_{r_i=1}^{L} \sum_{\delta=1}^{\frac{L-1}{2}} \frac{1}{\delta^\alpha} [J_1 (b_{r_i}^\dagger b_{r_i+\delta} + b_{r_i+\delta}^\dagger b_{r_i}) + \Delta (b_{r_i} b_{r_i+\delta} + b_{r_i+\delta}^\dagger b_{r_i}^\dagger)]. \tag{3.25}$$

Then, we can get

$$\hat{H} = \sum_k [b_k^\dagger b_k (J_0 + J_1 \Gamma_k^{(\alpha)}) + \frac{1}{2} \Delta \Gamma_k^{(\alpha)} (b_{-k} b_k + b_k^\dagger b_{-k}^\dagger)]. \tag{3.26}$$

Then, as $k \to 0$ and for $\alpha \notin \mathbb{N}_+$, the dispersion relation is:

$$E_k = \sqrt{(J_0 + J_1\Gamma_k^{(\alpha)})^2 - \Delta^2(\Gamma_k^{(\alpha)})^2} \tag{3.27}$$

$$\sim \left\{[(J_0 + J_1 a_1)^2 - \Delta^2 a_1^2] + 2[J_0 J_1 + (J_1^2 - \Delta^2)a_1]b_1 k^{\alpha-1} + (J_1^2 - \Delta^2)b_1^2 k^{2\alpha-2}\right. \tag{3.28}$$

$$\left. + 2[J_0 J_1 + (J_1^2 - \Delta^2)a_1]c_1 k^2 + 2(J_1^2 - \Delta^2)b_1 c_1 k^{\alpha+1} + (J_1^2 - \Delta^2)c_1^2 k^4\right\}^{1/2} \tag{3.29}$$

Therefore, the critical exponent α_c for generic fermionic quadratic Hamiltonians with long-range hoppings and pairings can be derived (Table 3.2).

The above results imply that fermionic quadratic Hamiltonians with long-range hoppings and pairings generally have $\alpha_c = 2$ except for a special case (no long-range hoppings) where $\alpha_c = \frac{3}{2}$. In comparison, bosonic quadratic Hamiltonians with long-range hoppings and pairings can have three possible values of α_c and especially, α_c can be 3 for gapless models. This notion of the maximum group velocity of quasiparticles can serve for the understanding of sublinear light cones in the research about information propagation in quantum systems.

3.7 Experimental relevance

Both the 1D and 2D long-range TFIMs have already been engineered in trapped-ion experiments for the purpose of quantum simulation, quantum computation, and quantum metrology [47, 48, 49, 50, 51]. Although the exact simulation of dynamics of large-scale quantum systems is generally hard on classical computers, our computation of quasiparticles

in long-range systems could provide classical benchmarks for the quality of these quantum simulators. For instance, as the quasiparticle energy in the confinement regime [108, 109] is also attainable from our RBM learning method, we can compare these energies to the oscillation frequency of magnetizations or other observables in experiments. Moreover, we can also compare the speed of correlation spreading in these quantum simulators with the maximal group velocity of quasiparticles directly derived from the obtained dispersion relations. These comparisons can provide guidance for the implementation of large-scale, and potentially high-dimensional, long-range quantum simulators and quantum computers.

3.8 Conclusion

We have used restricted Boltzmann machines (RBM) to learn the excited states of both 1D and 2D long-range interacting quantum spin models. We analyze the dispersion relations and identify the critical exponent where the maximal quasiparticles' group velocities transit from finite to divergent in the thermodynamic limit. In future research, it would be interesting to identify critical exponents for other Hamiltonians [110, 111, 112], study the effect of selecting different canonical spin configurations, and systematically compare the computational complexities of RBMs and tensor network states in high dimensions. It would also be interesting to investigate the relation between the critical α_c and Lieb-Robinson bounds [90].

Chapter 4: Deep neural networks as phase classifiers

4.1 Introduction

The application of NNs in representing quantum many-body states is a manifestation of the power of NNs as function approximators. NNs can also benefit the field of statistical physics by working as a data processor that can capture the intrinsic structure of data. Deep NNs have been widely used to identify phases and phase transitions based on their ability to detect multiple types of order parameters, including some nontrivial states without conventional order. It is shown that NNs can identify phase transitions in a range of correlated many-body systems just based on wave-function samples associated with ground states without the need to get the knowledge of locality conditions of Hamiltonians [3]. We are curious about how NNs behave in learning the phase transition points of quantum physical models with pairing terms. Such systems with particle-hole symmetry are often studied through analyzing their corresponding Bogoliubov-de-Gennes Hamiltonians. The energy eigenvalues of the eigenstates of the Bogoliubov-de-Gennes Hamiltonian can be positive, zero or negative, so there is no conventional ground state characterized by the smallest energy value for the NN to learn. We need to find other types of data that can be used to train the NNs to effectively capture the intrinsic differences between quantum states in distinct phases.

In this work [20], we study a 2D topologically nontrivial fermionic model Hamiltonian, analyze its phase diagram and demonstrate that deep NNs can learn the band-gap closing conditions only based on wave-function samples associated with several typical eigenstates (zero-energy edge modes), thus being able to identify the phase transition point without knowledge of Hamiltonians.

At first, we propose an effective model to create zero-energy edge modes that can be expressed in terms of Majorana operators at the edge of a honeycomb lattice by introducing textured pairings into a 2D topologically nontrivial Haldane model [113]. The key idea is to incorporate both the spin-singlet and textured spin-triplet pairings in the pseudospin-state dependent honeycomb optical lattice which breaks the TRS with complex next-nearest-neighbor (NNN) hopping. By tuning the pair coupling strength to match the amplitude and phase of hopping terms, zero-energy modes (also called zero modes) will arise on a single edge of the lattice. This is similar to the "sweet spot" conditions in Kitaev model. This suggests that to realize such zero modes, it is critical to break the 3-fold rotational symmetry of the pairing terms of the Hamiltonian, leading to a specific type of Majorana coupling. This requirement on the angular dependence of the sign of the spin-triplet pairing term is reminiscent of textured pairings in paired states of fermions [114].

In our model, the quantum system has a gapped SC phase and a gapless phase for parameters in the "sweet spot". Either phase can have a winding number $w = 0$ or $w \neq 0$. The phase diagram can be represented in the domain of phase parameters in the complex NNN hoppings. We demonstrate a method to reduce the gap-closing condition of the bulk Hamiltonian to the calculation of the discriminant. This circumvents the analytical complexity of a 4-band model. The value of this discriminant distinguishes the gapped

SC and gapless phases. It actually measures the "strength" of TRS breaking, thus further dividing the TRS-broken class into two groups. In the gapped SC phase, there always exist two pairs of zero modes while the winding number of bulk bands, $w = \pm 1$, is associated with extra normal gapless edge states. One of the zero-mode pairs can be fully pseudospin-polarized localized at an edge in special cases, while the other pair usually extends to deeper layers with exponentially decaying amplitudes. In the gapless phase, the second pair of zero modes vanishes due to their coupling with the bulk modes.

This work is organized as follows: In Sec. 4.2, we introduce our theoretical model Hamiltonian. In Sec. 4.3, we study the zero-energy edge modes from the aspects of the band structure, density profile and wavefunction symmetry by numerical simulation. In Sec. 4.4, the phase diagram of the quantum system is presented and discussed. At last, we describe a mathematical method to find a discriminant that gives the phase boundary between gapped SC and gapless phases. This discriminant characterizes the "strength" of TRS breaking.

4.2 Model and physical intuition

Our model is based on a generalized Haldane model in a pseudospin-state dependent honeycomb optical lattice [113], which is among many protocols proposed to implement the topological phases in systems of noninteracting fermions. In the realization of our system, ultracold atoms with two different hyperfine states would be described as two pseudospin states (spin-up "↑" and spin-down "↓"), each localized at one of two inequivalent sublattices (A and B). The natural tunneling between sites on the same sublattice and the

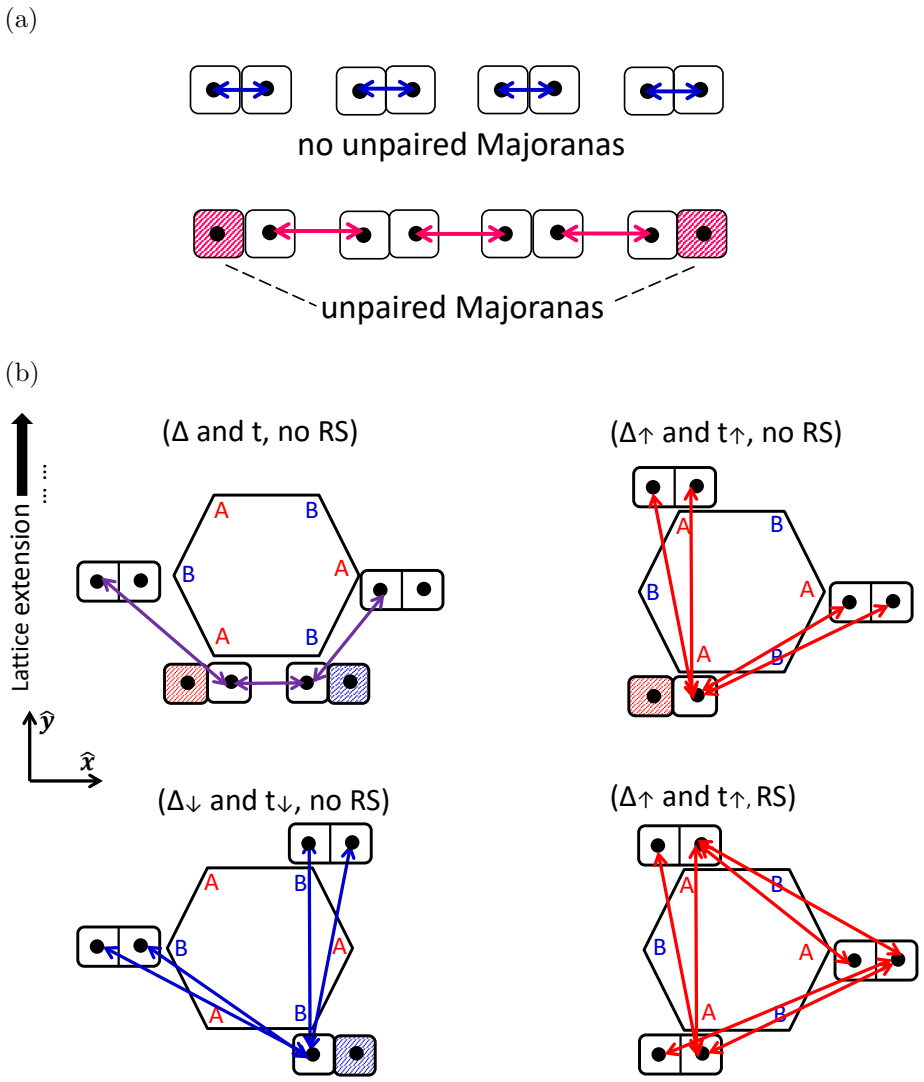


Figure 4.1: Physical intuition of generating zero-energy edge modes in the case of $\mu = 0$. (a) Kitaev's 1D spinless p-wave SC quantum wire. The two neighboring MFs constitute a normal fermion. The blue arrows in the upper chain signify the internal pairing of MFs with no unpaired MFs remaining. The red arrows in the lower chain indicate the inter-cell pairing of MFs with two unpaired MFs at the ends of the chain. (b) MF coupling at a single armchair edge of a 2D honeycomb lattice. The upper two and lower left subfigures are for our model in which there is no 3-fold rotational symmetry (RS). The solid bonds are the net Majorana couplings contributed by terms related to (Δ, t), $(\Delta_\uparrow, t_\uparrow)$ and $(\Delta_\downarrow, t_\downarrow)$ in $\hat{H}$, respectively. The shaded and colored cells denote the dangling MFs in a hexagon at a single armchair edge of the lattice. The lower right subfigure shows an example of unexpected MF couplings in which there is rotational symmetry, just in comparison with that in our model.

laser-induced coupling between different sublattices implement the next-nearest-neighbor (NNN) and nearest-neighbor hoppings, respectively. To generate unpaired MFs, we add spin-dependent pair interactions between atoms to obtain the total effective Hamiltonian:

$$\hat{H} = \hat{H}_0 + \hat{H}_p, \tag{4.1}$$

where

$$\begin{aligned}\hat{H}_0 = -t \sum_{\langle j,m \rangle} (a^\dagger_{\vec{r}_j} b_{\vec{r}_m} + h.c.) - t_\uparrow \sum_{\langle\langle j,j' \rangle\rangle} (e^{i\phi_A} a^\dagger_{\vec{r}_j} a_{\vec{r}_{j'}} + h.c.) \\ -t_\downarrow \sum_{\langle\langle m,m' \rangle\rangle} (e^{i\phi_B} b^\dagger_{\vec{r}_m} b_{\vec{r}_{m'}} + h.c.) \\ +\mu(\sum_j a^\dagger_{\vec{r}_j} a_{\vec{r}_j} - \sum_m b^\dagger_{\vec{r}_m} b_{\vec{r}_m}),\end{aligned} \tag{4.2}$$

$$\begin{aligned}\hat{H}_p = \sum_{\langle j,m \rangle} (\Delta a^\dagger_{\vec{r}_j} b^\dagger_{\vec{r}_m} + h.c.) + \sum_{\langle\langle j,j' \rangle\rangle, y_j < y_{j'}} (\Delta_\uparrow a^\dagger_{\vec{r}_j} a^\dagger_{\vec{r}_{j'}} + h.c.) \\ + \sum_{\langle\langle m,m' \rangle\rangle, y_m < y_{m'}} (\Delta_\downarrow b^\dagger_{\vec{r}_m} b^\dagger_{\vec{r}_{m'}} + h.c.).\end{aligned} \tag{4.3}$$

In the equations above, $\hat{H}_0$ is the original effective Hamiltonian following a unitary basis transformation as described in Ref. [113]; t is the nearest-neighbor hopping amplitude and $t_\uparrow(t_\downarrow)$ is the NNN hopping amplitude in sublattice $A(B)$. These three hopping amplitudes are real numbers. The NNN hopping phases are given by $\phi_A(\phi_B)$, and $\vec{r}_j(\vec{r}_m)$ denotes the site index of sublattice $A(B)$ (note that there is only one spin state at one site so that $\vec{r}_j$ and $\vec{r}_m$ are from different displacement vector sets). μ denotes half of the

chemical potential difference between sublattice A and B. $\hat{H}_p$ is the spin-dependent pairing Hamiltonian introduced by our model. Δ terms denote the pairings of atoms with opposite spins (spin-singlet), while $\Delta_\uparrow$ and $\Delta_\downarrow$ terms denote the pairings of atoms with the same spins (spin-triplet).

For simplicity, we first analyze a special case in which the on-site staggered potential vanishes, i.e., μ is zero. Then, the system can be viewed in the Majorana representation (Fig. 4.1(b)) by rewriting the Hamiltonian with Majorana operators:

$$a_{\vec{r}_j} = \frac{1}{2}(\gamma_{\vec{r}_j,\uparrow,1} + i\gamma_{\vec{r}_j,\uparrow,2}), b_{\vec{r}_m} = \frac{1}{2}(\gamma_{\vec{r}_m,\downarrow,1} + i\gamma_{\vec{r}_m,\downarrow,2}).$$

Here, $\{\gamma_{\vec{r}_j,\sigma,\alpha}, \gamma_{\vec{r}_m,\sigma',\beta}\} = 2\delta_{\vec{r}_j,\vec{r}_m}\delta_{\sigma\sigma'}\delta_{\alpha\beta}$ with each Majorana operator, γ, having three subscripts denoting the position, spin and Majorana type, respectively. In the following discussions, a Majorana coupling usually means a $i\gamma\gamma'$ term in the Hamiltonian and is denoted by a bond (double arrow) between two sites in Figs. 4.1(a) and 4.1(b). The Majorana coupling contributes an ordinary fermion which costs the energy of a bulk mode in the band structure. Based on the physical intuition shown in Fig. 4.1(b), we choose to introduce $\hat{H}_p$ in the above form, in order to cancel a part of the Majorana couplings introduced by the t, $t_\uparrow$ and $t_\downarrow$ hopping interactions. This choice yields the Hamiltonian in Appendix C.1.

The new "sweet spot" conditions that create dangling MFs at edges can be deduced. We inherit the key idea of Kitaev's 1D spinless p-wave SC chain model, which is to choose a specific type of Majorana coupling and leave unpaired MFs at the edges of the finite lattice (Fig. 4.1(a)). In our model, we choose to make the net effect of Δ and t terms be

the coupling $\gamma_{\vec{r}_j,\uparrow,2}\gamma_{\vec{r}_m,\downarrow,1}$ between nearest-neighbor sites, and the net effect of $\Delta_{\uparrow(\downarrow)}$ and $t_{\uparrow(\downarrow)}$ terms be the two couplings $\gamma_{\vec{r}_j,\uparrow(\downarrow),2(1)}\gamma_{\vec{r}_{j'},\uparrow(\downarrow),\alpha}$ between NNN sites with $y_j < y_{j'}$ and $\alpha = 1, 2$ (Fig. 4.1(b)), where y_j and $y_{j'}$ are the $\hat{y}$ component of $\vec{r}_j$ and $\vec{r}_{j'}$, respectively. It means that the coupling connecting A and B sublattices is only between type-2 MFs in A and type-1 MFs in B. And the coupling within $A(B)$ sublattice is only from type-2(1) MFs to MFs of both two types with bigger indices in $\hat{y}$ direction. These requirements can be reduced to Eqs. (4.4)-(4.6) which can be called the "sweet spot" in the parameter space. So the final result is that the type-1 MFs of atoms in A sublattice and type-2 MFs of atoms in B sublattice at one armchair edge (the shaded and colored cells in Fig. 4.1(b)) are isolated, i.e., $\gamma_{y_1,\uparrow,1}$ and $\gamma_{y_1,\downarrow,2}$ do not appear in $\hat{H}$, as there are no other MFs outside the lattice to couple with them.

$$\Delta = -t, \tag{4.4}$$

$$\Delta_\uparrow = -t_\uparrow e^{-i\phi_A}, \tag{4.5}$$

$$\Delta_\downarrow = t_\downarrow e^{-i\phi_B}. \tag{4.6}$$

Therefore, our model has generalized "sweet spot" conditions analogous to those of a 1D Kitaev chain, and actually possesses textured pairings analogous to that in the original model of fermionic paired states [114]. Note first that the $\Delta_{\uparrow(\downarrow)}$ terms are written in a particular order ($y_{j(m)} < y_{j'(m')}$) to avoid mixing definitions. Second, it is critical to break the 3-fold rotational symmetry of the net Majorana couplings within a hexagon

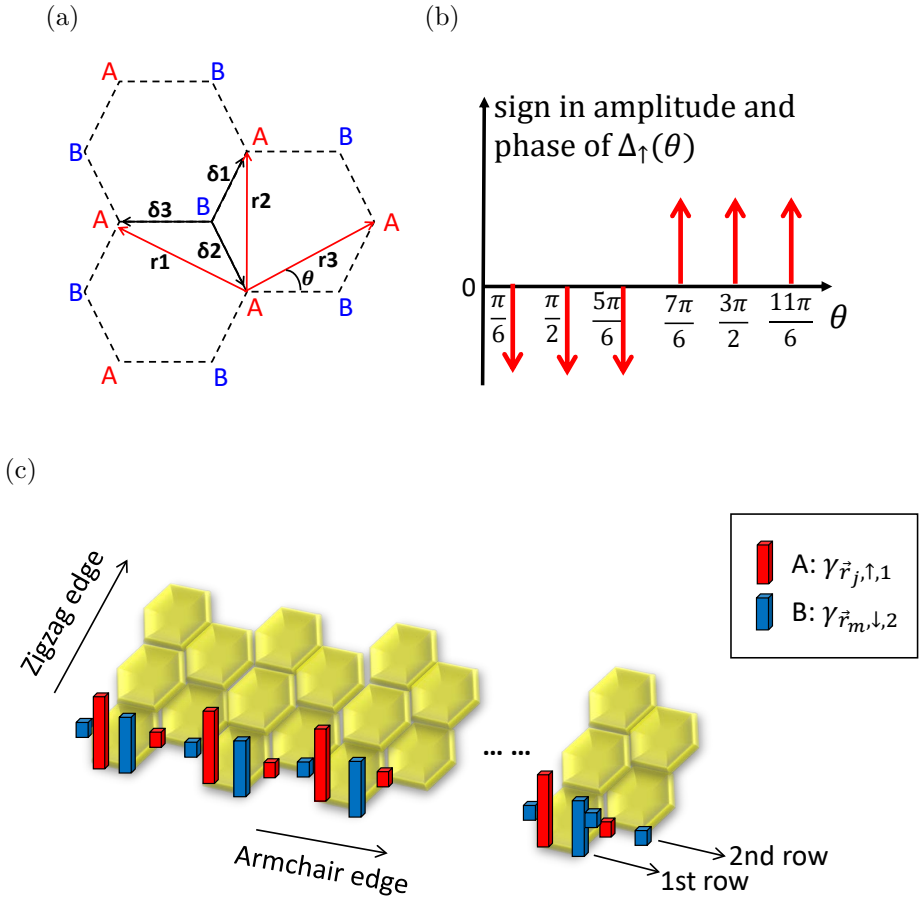


Figure 4.2: Schematic of our physical system. (a) The depictions of the three vectors ($\vec{\delta}_j$) connecting nearest-neighbor sites and the three vectors ($\vec{r}_j$) connecting NNN sites for $j = 1, 2, 3$. (b) Angular distribution of the sign in the amplitude and phase of the defined pairing which is similar to the domain wall structure. (c) The distribution of zero modes at a single armchair edge. The heights of the pillars denote the amplitudes of the wave function at each site.

(Fig. 4.1(b)). Since t is real, the net coupling between nearest-neighbor A and B sites may be reduced to one bond (the upper left subfigure in Fig. 4.1(b)). However, there is no degree of freedom to reduce the net coupling between NNN A or B sites to less than 2 bonds since t_σ is complex. Thus, any pairing Hamiltonian that has 3-fold rotational symmetry (such as the lower right subfigure in Fig. 4.1(b)) does not allow dangling MFs at edges. We need to impose requirements on the sign in the amplitude and phase of Δ_σ to break the 3-fold rotational symmetry.

We present the requirement on pairing terms in an alternative way, which highlights the concept of textures in pairing. Following the reference [113], choose the Peierls phase associated with the NNN hopping path $a^\dagger_{\vec{r}_j} a_{\vec{r}_{j'}}$ to be:

$$\phi_A(j, j') = -\vec{p} \cdot (\vec{r}_j - \vec{r}_{j'})/2.$$

A pair creation term in $\hat{H}$ is $\Delta_\uparrow(\theta) a^\dagger_{\vec{r}_j} a^\dagger_{\vec{r}_{j'}}$, where θ is the angle between $(\vec{r}_{j'} - \vec{r}_j)$ and the $\hat{x}$ axis. Then, we can deduce from the above requirements:

$$\Delta_\uparrow(\theta) = \begin{cases} -t_\uparrow e^{-i\vec{p}\cdot(\vec{r}_{j'} - \vec{r}_j)/2}, & (y_j < y_{j'}) \\ t_\uparrow e^{i\vec{p}\cdot(\vec{r}_{j'} - \vec{r}_j)/2}. & (y_j > y_{j'}) \end{cases}$$

We can see that the sign in $\Delta_\uparrow(\theta)$ in front of the amplitude $t_\uparrow$ and phase $\vec{p}\cdot(\vec{r}_{j'} - \vec{r}_j)/2$ is changed across $[0, 2\pi]$ (Fig. 4.2(b))). This requirement of the broken rotational symmetry leads to exotic textures in the pairing terms [114], and the angular distribution of the sign in the spin-triplet pairing term has a reorientation similar to that of the domain wall structure in magnetism [115].

4.3 Zero-energy edge modes

The geometric definition of our model is depicted in Fig. 4.2(a). We use three displacement vectors $\vec{\delta}$ to denote the nearest-neighbor hoppings and another three $\vec{r}$ to denote the NNN hoppings:

$$\vec{\delta}_1 = a(\frac{1}{2}, \frac{\sqrt{3}}{2}), \vec{\delta}_2 = a(\frac{1}{2}, -\frac{\sqrt{3}}{2}), \vec{\delta}_3 = a(-1, 0),$$
$$\vec{r}_1 = a(-\frac{3}{2}, \frac{\sqrt{3}}{2}), \vec{r}_2 = a(0, \sqrt{3}), \vec{r}_3 = a(\frac{3}{2}, \frac{\sqrt{3}}{2}).$$

Here, a is the side length of a hexagon plaquette and is set as the unit of length in this work. Using conventions of current techniques of implementing the complex NNN hopping by laser-induced transitions [113], we choose the Peierls phase associated with $a^\dagger_{\vec{r}_j} a_{\vec{r}_{j'}}$ to be $\phi_A(j, j') = -\vec{p} \cdot (\vec{r}_j - \vec{r}_{j'})/2$ and similarly $\phi_B(m, m') = \vec{p} \cdot (\vec{r}_m - \vec{r}_{m'})/2$, where $\vec{p}$ is the momentum transfer associated with the laser-induced tunneling.

We calculate the band structure using a momentum space representation based on a Fourier transformation in the $\hat{x}$ direction:

$$\begin{aligned} \hat{a}_{\vec{r}_j} &= \frac{1}{\sqrt{N_x}} \sum_{k_x} e^{ik_x x_j} \hat{a}_{k_x, y_j}, \\ \hat{b}_{\vec{r}_m} &= \frac{1}{\sqrt{N_x}} \sum_{k_x} e^{ik_x x_m} \hat{b}_{k_x, y_m}. \end{aligned}$$

$x_{j(m)}$ and $y_{j(m)}$ are the components of $\vec{r}_{j(m)}$ in x and y direction, respectively. N_x is the number of cells along $\hat{x}$ direction, much larger than that along $\hat{y}$ direction. Thus the

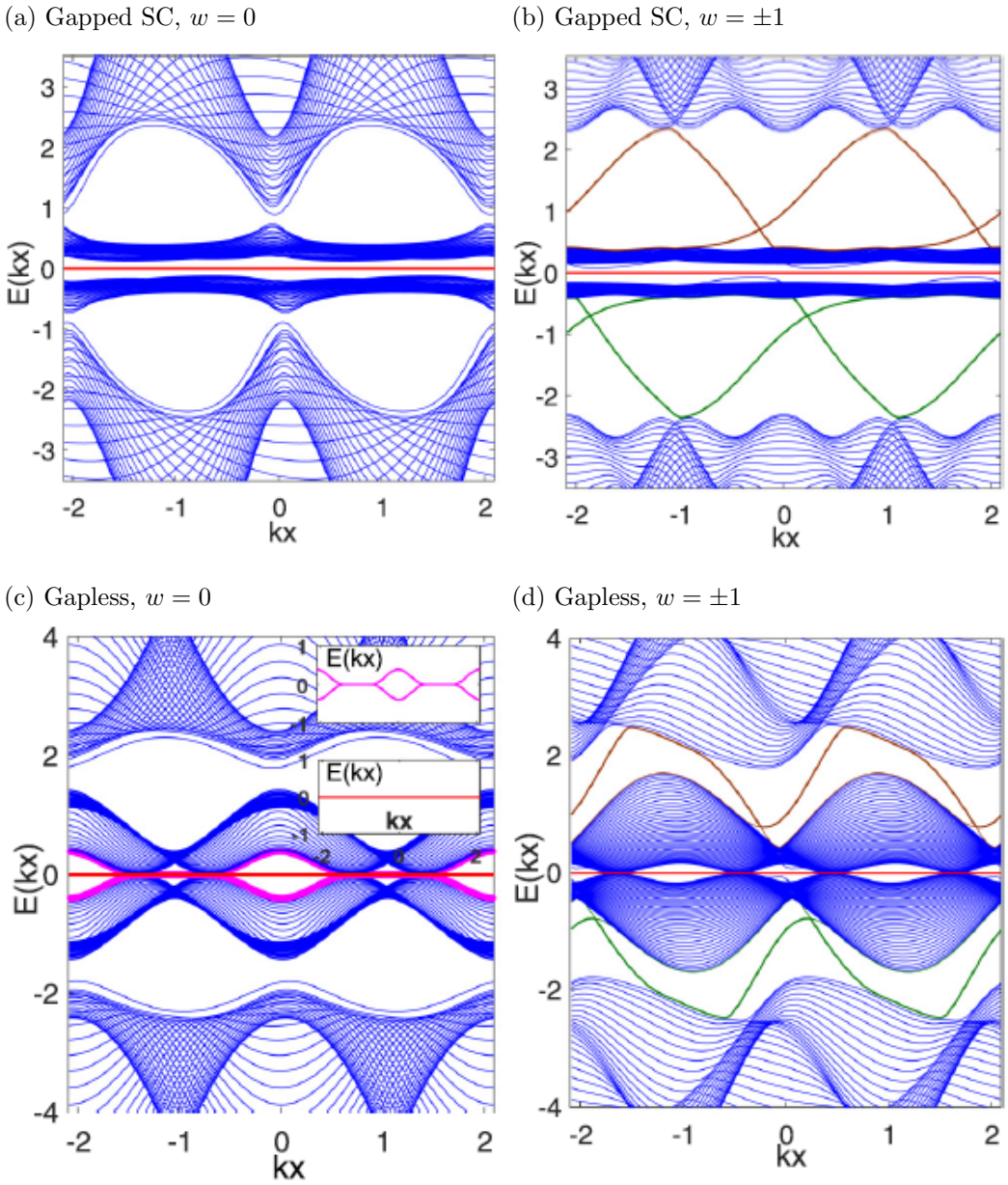


Figure 4.3: Band structure for $\mu = 0$ in four cases characterized by different gap conditions and winding numbers w. Fixed parameters: $t = 1$, $t_\uparrow = 0.4$, $t_\downarrow = 0.6$. (a) $\vec{p} = (0.9K_x, 0)$; (b) $\vec{p} = (0, 3K_y)$; (c) $\vec{p} = (0.1K_x, 0.2K_y)$; (d) $\vec{p} = (0.6K_x, 3K_y)$, where $K_x = 2\pi/3$ and $K_y = K_x/\sqrt{3}$. Each group of blue curves represents a bulk band. The dark red straight lines are the zero-energy edge modes and the brown and green curves in (b) and (d) are gapless edge modes. In (a) and (b), the zero modes are 4-fold degenerate; in (c) and (d), the zero modes are 2-fold degenerate. The inset in (c) zooms in on the split zero modes (magenta) and completely flat zero modes (dark red) separately.

basis vector of the Bogoliubov-de-Gennes Hamiltonian of $\hat{H}$ for particular k_x becomes $(\hat{a}_{k_x,y_1}, ..., \hat{b}_{k_x,y_1}, ..., \hat{a}^\dagger_{-k_x,y_1}, ..., \hat{b}^\dagger_{-k_x,y_1}, ...)^T$, in which the subscripts denoting y component range over all the rows (also called layers in this work). Then, the band structure containing both the edge modes and bulk bands is obtained by the diagonalization of the Hamiltonian in this basis. The results are shown in Figs. 4.3(a)–4.3(d).

Our model has a gapless and a gapped SC phase at the "sweet spot", either of which may have zero or nonzero winding numbers of the first excited band. As a common feature, there are two groups of bulk bands (blue curves) in each of the upper or lower half-planes. These are inherited from the Haldane model due to the number of inequivalent sites in a unit cell. The gapped SC phase and gapless phase are distinguished by the gap closing condition between the zero-energy line and the first excited band in the upper half-plane ("band 1"). In the gapped SC phase (Figs. 4.3(a) and 4.3(b)), there are two pairs of zero modes with complete flat bands. These are shown by straight red lines in Figs. 4.3(a)–4.3(d), coinciding with each other. By contrast, the gapless phase holds one pair of zero modes while the other pair of zero modes partially merges into the bulk modes in some ranges of k_x (Figs. 4.3(c) and 4.3(d)). Furthermore, each of the gapped SC and gapless phases can have winding numbers $w = 0$ or $w = \pm 1$ of the band 1. The nonzero-winding-number phase corresponds to additional ordinary gapless edge modes between the band 1 and the second excited bulk band in the upper half-plane (band 2).

To further support the correctness of our model, we get the wavefunction of the zero modes in the band structure as below by numerical simulation:

$$
\begin{aligned}
\hat{\Psi}_{k_x} &= \sum_{y_j} (u_{k_x,y_j}, v_{k_x,y_j}, u_{k_x,y_j}, -v_{k_x,y_j}) \\
&\quad \times (\hat{a}_{k_x,y_j}, \hat{b}_{k_x,y_j}, \hat{a}^{\dagger}_{-k_x,y_j}, \hat{b}^{\dagger}_{-k_x,y_j})^T \\
&\approx \sum_{x_j} (\frac{1}{\sqrt{N_x}} u_{k_x,y_1} e^{-ik_x x_j}) \hat{\gamma}^{(A)}_{x_j,y_1,1} \\
&+ \sum_{x_m} (\frac{1}{\sqrt{N_x}} v_{k_x,y_1} e^{-ik_x x_m}) \hat{\gamma}^{(B)}_{x_m,y_1,2},
\end{aligned} \tag{4.7}
$$

where u_{k_x,y_i} and v_{k_x,y_i} are the wavefunctions for the ith layer in the $\hat{y}$ direction using the partial Fourier-transformed basis. Note that there is a degree of freedom in choosing the coefficient in front of a specific eigenstate. We use the approximation sign and just keep the wavefunctions for the first layer in the above as the numerical simulation indicates that the solution amplitudes generally decay at an exponential rate (Fig. 4.2(c)).

4.4 Phase diagram at the sweet spot

The phase diagram of the quantum system at the "sweet spot" is worth analyzing. For fixed amplitude parameters (t, $t_\uparrow$ and $t_\downarrow$), the phase of the system varies with the NNN hopping phases ϕ_A and ϕ_B. This is displayed by the phase diagram in the momentum coordinates p_x and p_y, which is shown in Figs. 4.4(a) and 4.4(b). There are a total of four phases associated with the two alternatives "gapped SC vs. gapless" and "$w = 0$ vs. $w = \pm 1$".

As we illustrated in Sec. 4.3, the phase boundary between the gapped SC phase and gapless phase can be deduced by the bulk-edge correspondence in topological physics. In

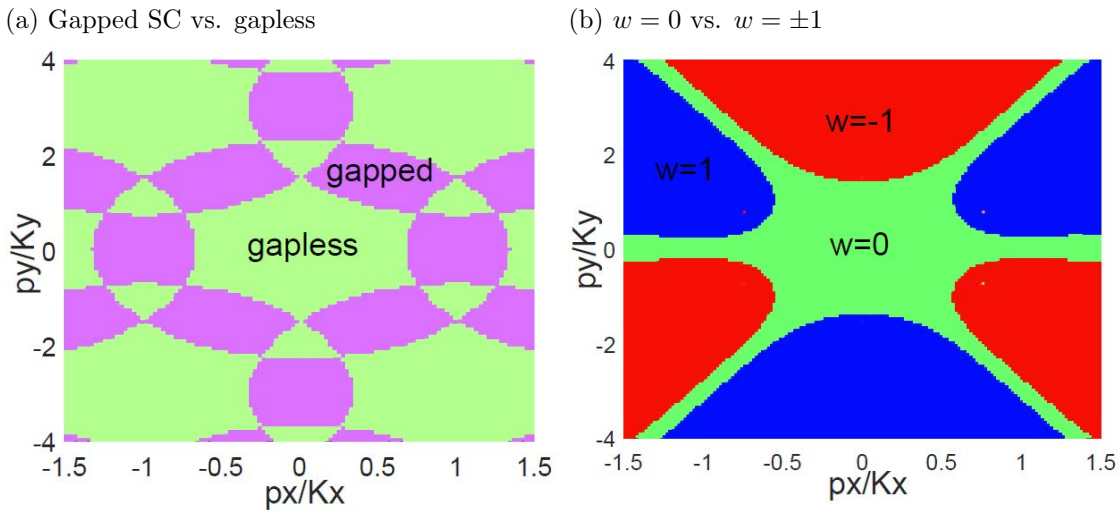


Figure 4.4: Phase diagram at the generalized "sweet spot" when $\mu = 0$, obtained by numerical simulation. (a) shows the phase boundary between the gapped SC phase (purple) and gapless phase (light green). (b) shows the phase boundary between phases with different winding numbers. The green region indicates $w = 0$, the blue region $w = 1$ and the red region $w = -1$. A combination of these two figures shows the full 4-phase diagram.

the above version of our model, the pseudospin space and the particle-hole space each contribute 2 degrees of freedom. Thus, we have a 4-band model, the description of which requires solutions of a quartic equation, which in general have complicated analytical forms. To circumvent the mathematical complexity of analytical band expressions of a 4-band model, we demonstrate a method to rigorously reduce into a discriminant the parameter conditions of the gap closing between band 1 and band 2 identified in Sec. 4.3.

In the gapped SC phase, the gap between bands 1 and 2 is open due to TRS breaking by complex NNN hoppings, protecting the zero modes from coupling with bulk modes. In a certain range, this gap is approximately proportional to the amplitude of the relative complex hopping $|\frac{t_\sigma}{t}|$. At the "sweet spot", the Bogoliubov-de-Gennes Hamiltonian in momentum space, $H_{BdG}(\vec{k})$, reduces to:

$$\begin{pmatrix} -2t_\uparrow f_- + \mu & -tg^* & |\Delta_\uparrow|h & \Delta g^* \\ -tg & -2t_\downarrow f_+ - \mu & -\Delta g & |\Delta_\downarrow|h^* \\ |\Delta_\uparrow|h^* & -\Delta g^* & 2t_\uparrow f_+ - \mu & tg^* \\ \Delta g & |\Delta_\downarrow|h & tg & 2t_\downarrow f_- + \mu \end{pmatrix},$$

where

$$\begin{aligned} f_+ &= f_+(\vec{k},\vec{p}) = \sum_{j=1}^{3} \cos((\vec{k}+\vec{p}/2)\cdot\vec{r_j}), \\ f_- &= f_-(\vec{k},\vec{p}) = \sum_{j=1}^{3} \cos((\vec{k}-\vec{p}/2)\cdot\vec{r_j}), \\ g &= g(\vec{k}) = \sum_{j=1}^{3} e^{-i\vec{k}\cdot\vec{\delta_j}}, \\ h &= h(\vec{k},\vec{p}) = 2i\sum_{j=1}^{3} e^{-i\frac{\vec{p}}{2}\cdot\vec{r_j}} \sin(\vec{k}\cdot\vec{r_j}). \end{aligned}$$

For a general quartic equation in E, $F(E) = \prod_{j=1}^{4}(E - E_j) = 0$, a two-fold root $E = 0$, corresponding to band-touching conditions, implies the constraint $F(E=0) = \frac{dF}{dE}(E=0) = 0$. The energy eigenvalues, E, can be calculated by solving the characteristic polynomial $F(E) = \det(H_{BdG}(\vec{k}) - E * I_{4\times4}) = 0$, where det() means the determinant. By invoking the "sweet spot" conditions Eqs. (4.4)–(4.6), $\mu = 0$ and the band constraint, we get:

$$F(E=0) = (t_\uparrow t_\downarrow)^2(4f_+f_- + |h|^2)^2, \tag{4.8}$$

$$\frac{dF}{dE}(E=0) = 2t_\uparrow t_\downarrow(t_\uparrow - t_\downarrow)(f_+ - f_-)(4f_+f_- + |h|^2). \tag{4.9}$$

We make three observations that lead to a fuller understanding of the solutions of the Eqs. (4.8)–(4.9). First, these equations contain a common nonnegative factor (discriminant)

$$4f_+f_- + |h|^2 = 4|\sum_{j=1}^{3} e^{i\vec{k}\cdot\vec{r}_j}\cos(\vec{p}\cdot\vec{r}_j/2)|^2. \tag{4.10}$$

Thus, when Eq. (4.8) vanishes ($t_\uparrow t_\downarrow \neq 0$ typically), Eq. (4.9) also vanishes. So there must be 0 or 2 bands (or more bands) simultaneously touching the zero-energy line. This is consistent with the fact that the two intermediate bulk bands touch with the zero-energy line at the same points in the band structure.

Second, the terms t and $g(\vec{k})$ do not affect the band touching condition. Thus, for our particular MF coupling, there is no net effect of nearest-neighbor interactions on the phase at the "sweet spot".

Third, the effects of geometry and of parameter strengths are independent. Since $t_\uparrow$ and $t_\downarrow$ are nonzero and unequal in most cases, we just need to focus on the equality $4f_+f_- + |h|^2 = 0$. As shown in the appendix, this is equivalent to:

$$|\cos(\vec{p}\cdot\vec{r}_j/2)| + |\cos(\vec{p}\cdot\vec{r}_k/2)| \geq |\cos(\vec{p}\cdot\vec{r}_l/2)|, \tag{4.11}$$

for any (j,k,l) being a permutation of $(1,2,3)$.

The above inequalities describe the parameter range of the gapless phase as compared to the gapped SC phase and actually give a measure of the "strength" of TRS breaking.

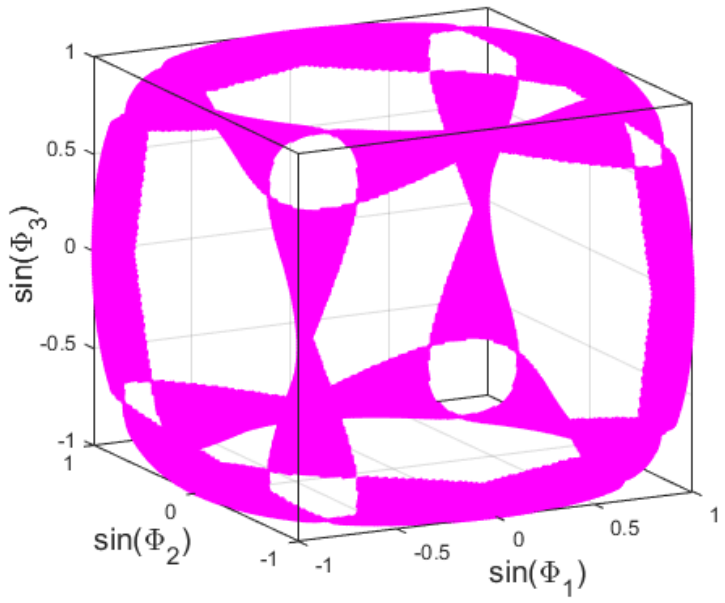


Figure 4.5: Regions of the gapped SC phase in the 3D parameter space ($\sin\Phi_1$, $\sin\Phi_2$, $\sin\Phi_3$). The shaded surface near the 12 edges of the cube shows the domain in which at least one inequality $|\cos\Phi_j|+|\cos\Phi_k| \geq |\cos\Phi_l|$ is violated, where (j,k,l) is a permutation of $(1,2,3)$. This surface defines the gapped SC phase. Note that the only points in this parameter space that have physical significance are those for which $\Phi_2 = \Phi_1+\Phi_3$ (including both the near-edge and kernel regions).

The phase diagram in Fig. 4.4(a) obtained by numerical simulations of band structures at every point in $p_x - p_y$ plane is exactly the same with that obtained according to the three inequalities. Defining the Peierls phases associated with the complex NNN hoppings inside a hexagon as $\vec{p} \cdot \vec{r}_j/2 = \Phi_j$ (with restriction $\Phi_2 = \Phi_1 + \Phi_3$), the distribution of gapped SC phase in the $(\sin\Phi_1, \sin\Phi_2, \sin\Phi_3)$ parameter space is shown in Fig. 4.5. It can be seen that the gapped SC phase is mainly localized near the edge of the cube, which indicates that at least two of the three $|\sin\Phi_j|$ are near 1. We know that $2t_\sigma \sin\Phi_j$ is the amplitude difference of NNN hopping terms before and after the TRS transformation. So we claim that the gapped SC phase corresponds to "strong" TRS breaking in which at least one of the three inequalities (4.11) is violated. For "strong" TRS breaking, two of the three Peierls phases lead to a relatively large energy difference after the TRS transformation, and the energy gap is large enough to protect the zero modes from coupling with bulk modes. By contrast, the gapless phase corresponds to "weak" TRS breaking, which is supported by its occupying the central part near $(p_x = 0, p_y = 0)$ in the phase diagram. Thus, this measure further divides the TRS-broken class into two groups.

Moreover, the phase diagram showing "$w = 0$ vs. $w = \pm 1$" (Fig. 4.4(b)) has 3-fold rotational symmetry. This phase diagram is obtained by numerically calculating [116] the winding number, w, for every point in $p_x - p_y$ plane:

$$w = \frac{i}{2\pi} \int_{T^2} (\langle \partial_{k_x} u_1 | \partial_{k_y} u_1 \rangle - \langle \partial_{k_y} u_1 | \partial_{k_x} u_1 \rangle) d^2k, \tag{4.12}$$

where $|u_1\rangle$ is the eigenstate of band 1 and T^2 is the first Brillouin zone. Since the sum

of the topological invariants of the neighboring band 1 and band 2 are zero, it is enough to do such a calculation for one band. In the pattern of the phase diagram, the center is mainly the topologically trivial region, $w = 0$, while the surrounding $w = 1$ (blue) and $w = -1$ (red) nontrivial regions are separated by "thin-ribbon" regions with $w = 0$. The central lines of the six "thin-ribbon" green regions indicate $\vec{p} \perp \vec{r}_j$, for which one of the three Peierls phases is zero. The topological transition only occurs when the gap between band 1 and band 2 closes so that the topology of bulk bands intrinsically changes. The bulk bands only touch at the typical Dirac points $(K_x, \pm K_y)$ ($K_x = \frac{2\pi}{3}$ and $K_y = \frac{2\pi}{3\sqrt{3}}$), which indicates the decoupling of the two sublattices. Band touching at only one of the two inequivalent Dirac points entails the change of winding number by 1, while the case of touching at both has not been observed in this model. Additionally, when w is nonzero, the system is characterized by ordinary gapless edge modes connecting band 1 and band 2 at two armchair edges.

In general, when $\mu \neq 0$, the system displays changes in both of the gapped SC and gapless phases. When $|\mu|$ increases from zero in the gapped SC phase, the gap between band 1 and the zero-energy line gradually decreases until two zero modes couple with the bulk modes, leading to a decreasing jumping of the degeneracy of zero modes. The pair of fully pseudospin-polarized zero modes that is exactly localized at the first layer in $\mu = 0$ case, will extend to the deeper layers with exponentially decreasing amplitudes. In the gapless phase, the pairing occurs in a partially filled band. The Dirac points connecting band 1 and band 2 vanish, accompanied by the relative displacement of the upper and lower Dirac cones. The gapless property is preserved but the bands are indirectly closed.

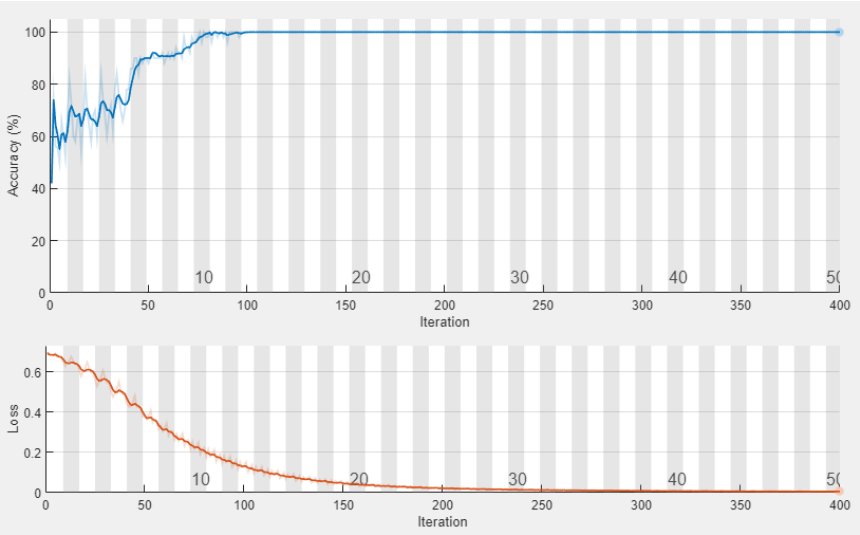
At last, it should be mentioned that our theoretical model seems to be ahead of

the development of experimental techniques in fermionic optical lattices. Some of the necessary techniques have been well developed as described in the introduction. The biggest challenge may be the precise tuning of both the spin-singlet and spin-triplet pairings. Recent progress in implementing pairings in cold atoms includes finite-momentum Cooper pairings. These can induce Fulde-Ferrell (FF) states which have superconducting order parameters with uniform amplitudes but spatially dependent phases. Protocols for creating such FF superfluids in spinful [117, 118, 119, 120, 121] and spinless [122] cold atoms may be helpful for implementing the crucial Hamiltonian components in our model, though it is still a big step. Further experimental progress is needed to realize our model, but our study of this model Hamiltonian may be insightful for researchers in cold atom physics.

4.5 Application of neural networks as phase classifiers

After fully understanding the phase diagram of the above model, we use a deep NN to learn the phase transition of the model between the gapped SC phase and the gapless phase with a winding number of 0. We treat it as a supervised ML task. The training data used to feed the NN is a collection of composite column vectors, each of which is formed by aligning the four intermediate eigenvectors (whose eigen-energy values center around the zero-energy) associated with $k_x = 0$ in a single column. With a fixed parameter $p_y = 0$ and p_x varying over a small range far away from the phase-transition point $p_x = 2K_x/3$ that can be obtained through theoretical analysis, we can get a set of such composite wave-function vector samples, each of which is associated with a known phase as the so-called label. After a full training of the NN, we generate the test data which is a collection of

(a)



(b)

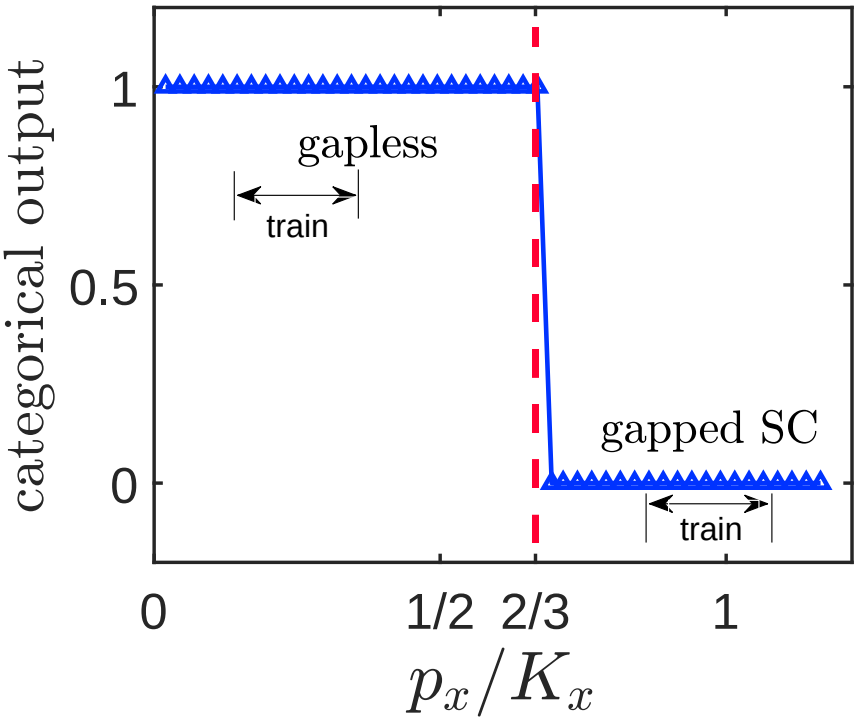


Figure 4.6: (a) Training process using the deep-learning toolbox provided by Matlab software. The prediction accuracy increases and the loss decreases as the training proceeds. (b) Effects of learning the phase transition between the gapped SC phase and the gapless phase (both with a winding number of 0). p_y/K_y is fixed to be 0. The exact phase transition point is $p_x/K_x = 2/3$. The output of a fully trained deep NN exhibits a sudden change when p_x/K_x varies across the phase transition point. The wave-function samples used as training data just cover a small parameter range far away from the critical point.

the composite wave-function vectors with p_x covering the full interval from 0 to $1.2K_x$ and check its accuracy in phase prediction. It turns out that the NN can return an output which exhibits a sudden change when p_x varies across the phase transition point, thus being able to perform a phase classification (Fig. 4.6(b)). We argue that the success of NNs in phase classification for the above model Hamiltonian is possibly based on that NNs can learn the features about whether all the four intermediate eigenstates are edge modes with zero energy (corresponding to the gapped SC phase) or one pair of the four intermediate eigenstates couple with bulk modes and thus are split zero modes (corresponding to the gapless phase). Therefore, this work justifies that deep NNs can learn the band-gap closing conditions only based on wave-function samples associated with several typical eigenstates (zero-energy edge modes), thus being able to identify a specific phase-transition point without knowledge of Hamiltonians. The training process is shown in Fig. 4.6(a).

4.6 Conclusion

We have proposed a method for creating zero-energy edge modes of a honeycomb optical lattice of ultracold atoms. This is done by generalizing a 2D topologically nontrivial Haldane model and introducing textured pairings. Both the spin-singlet and textured spin-triplet pairings are added to a pseudospin-state dependent lattice, whose time-reversal symmetry is broken by complex next-nearest-neighbor hoppings. If generalized "sweet spot" conditions are satisfied, zero modes will arise on a single edge of the lattice. We have analyzed their properties, such as pseudospin polarization. We find that this system has a gapped superconducting phase and a gapless phase, each of which can have zero

or nonzero winding numbers, and have calculated the phase diagrams of the system. We have simplified the understanding of the bandgap-closing condition of the bulk Hamiltonian by identifying a discriminant that distinguishes the gapped superconducting and gapless phases and provides a measure of the "strength" of time-reversal symmetry breaking. Finally, we investigate the application of deep neural networks in phase classification without knowledge of model Hamiltonians.

Appendix A: Appendix for Chapter "Efficiency of neural-network state representations of one-dimensional quantum spin systems"

A.1 Proof of the convergence of $|\Psi^{(L,\infty)}\rangle$ for long-range-fast-decay RBMs

In this section, we prove the convergence of $|\Psi^{(L,\infty)}\rangle$ which satisfies the two conditions in the definition of the long-range-fast-decay (LRFD) RBMs given in Sec. 2.2.1.

We have defined

$$\psi^{(L,\infty)}(\vec{\sigma}) = \prod_{j=1}^{L} e^{a_j^{(L)}\sigma_j} \prod_{k=1}^{\infty} \cosh(b_k^{(L)} + \sum_{j=1}^{L} \sigma_j W_{j,k}^{(L)}). \tag{A.1}$$

Its corresponding truncated state sequence is: $|\Psi^{(L,N_h)}\rangle = \sum_{\vec{\sigma}} \psi^{(L,N_h)}(\vec{\sigma})|\vec{\sigma}\rangle$, where

$$\psi^{(L,N_h)}(\vec{\sigma}) = \prod_{j=1}^{L} e^{a_j^{(L)}\sigma_j} \prod_{k=1}^{N_h} \cosh(b_k^{(L)} + \sum_{j=1}^{L} \sigma_j W_{j,k}^{(L)}). \tag{A.2}$$

For any fixed L and any $\vec{\sigma}$, define

$$A_n(\vec{\sigma}) = |\psi^{(L,nL)}(\vec{\sigma})|, \tag{A.3}$$

$$\phi_n(\vec{\sigma}) = \arg(\psi^{(L,nL)}(\vec{\sigma})). \tag{A.4}$$

Then

$$\psi^{(L,nL)}(\vec{\sigma}) = A_n(\vec{\sigma})e^{i\phi_n(\vec{\sigma})}. \tag{A.5}$$

In the following part, we will prove that the sequence $\{\psi^{(L,nL)}(\vec{\sigma}) : n \in \mathbb{N}\}$ is a Cauchy sequence in $\mathbb{C}$ with d as a metric, where $d : \mathbb{C} \times \mathbb{C} \to \mathbb{R}$ is just the commonly used distance between two complex numbers defined as $d(z_1, z_2) = |z_1 - z_2|$ for $z_1,\ z_2 \in \mathbb{C}$.

Proof. If there exists any finite $n' \in \mathbb{N}$ such that $\psi^{(L,n'L)}(\vec{\sigma}) = 0$, then $\psi^{(L,nL)}(\vec{\sigma}) = 0$ for all $n \geq n'$. Then $\{\psi^{(L,nL)}(\vec{\sigma}) : n \in \mathbb{N}\}$ must be a Cauchy sequence in the metric space $(\mathbb{C}, d)$. So we will just focus on the cases where $\psi^{(L,nL)}(\vec{\sigma}) \neq 0$, which implies $A_n(\vec{\sigma}) \neq 0$, for all $n \in \mathbb{N}$ in the following part.

For any $n > \tilde{k}_s, n \in \mathbb{N}$,

$$A_n(\vec{\sigma}) = \Big| \prod_{j=1}^{L} e^{a_j^{(L)}\sigma_j} \prod_{k=1}^{nL} \cosh(b_k^{(L)} + \sum_{j=1}^{L} \sigma_j W_{j,k}^{(L)}) \Big| \tag{A.6}$$

$$= \Big| \prod_{j=1}^{L} e^{a_j^{(L)}\sigma_j} \prod_{k=1}^{\tilde{k}_s L} \cosh(\theta_k(\vec{\sigma})) \Big| \Big| \prod_{k=\tilde{k}_s L+1}^{nL} \cosh(\theta_k(\vec{\sigma})) \Big| \tag{A.7}$$

$$= A_0 \Big| \prod_{k=\tilde{k}_s L+1}^{nL} \cosh(\theta_k(\vec{\sigma})) \Big|. \tag{A.8}$$

Define

$$u_k(\vec{\sigma}) = \mathrm{Re}(b_k^{(L)} + \sum_{j=1}^{L} \sigma_j W_{j,k}^{(L)}) \in \mathbb{R}, \tag{A.9}$$

$$v_k(\vec{\sigma}) = \mathrm{Im}(b_k^{(L)} + \sum_{j=1}^{L} \sigma_j W_{j,k}^{(L)}) \in \mathbb{R}, \tag{A.10}$$

$$U_k = \max_{\{\vec{\sigma}\}}\{|u_k(\vec{\sigma})|\} = |\,\mathrm{Re}(b_k^{(L)})| + \sum_{j=1}^{L} |\,\mathrm{Re}(W_{j,k}^{(L)})|, \tag{A.11}$$

$$V_k = \max_{\{\vec{\sigma}\}}\{|v_k(\vec{\sigma})|\} = |\,\mathrm{Im}(b_k^{(L)})| + \sum_{j=1}^{L} |\,\mathrm{Im}(W_{j,k}^{(L)})|. \tag{A.12}$$

$$A_n(\vec{\sigma}) = A_0 \prod_{k=\tilde{k}_s L+1}^{nL} \Big|\cosh(u_k(\vec{\sigma}) + i v_k(\vec{\sigma}))\Big| \tag{A.13}$$

$$= A_0 \prod_{k=\tilde{k}_s L+1}^{nL} \frac{1}{2}\sqrt{e^{2u_k(\vec{\sigma})} + e^{-2u_k(\vec{\sigma})} + 2\cos(2v_k(\vec{\sigma}))}. \tag{A.14}$$

So we can upper bound $A_n(\vec{\sigma})$ by

$$A_n(\vec{\sigma}) \le A_0 \prod_{k=\tilde{k}_s L+1}^{nL} \frac{1}{2}\sqrt{e^{2u_k(\vec{\sigma})} + e^{-2u_k(\vec{\sigma})} + 2} \tag{A.15}$$

$$\le A_0 \prod_{k=\tilde{k}_s L+1}^{nL} \exp(\frac{1}{2}u_k^2(\vec{\sigma})) \tag{A.16}$$

$$\le A_0 \prod_{k=\tilde{k}_s L+1}^{nL} \exp(\frac{1}{2}U_k^2) \tag{A.17}$$

$$= A_0 \exp\Big(\frac{1}{2}\sum_{k=\tilde{k}_s L+1}^{nL} (|\,\mathrm{Re}(b_k^{(L)})| + \sum_{j=1}^{L} |\,\mathrm{Re}(W_{j,k}^{(L)})|)^2\Big). \tag{A.18}$$

Using the upper bounds in the definition of LRFD RBMs, we can get

$$A_n(\vec{\sigma}) \leq A_0 \exp\Big[\frac{1}{2}\sum_{k=\tilde{k}_s L+1}^{nL}\Big(\lambda_R(\tilde{k}(k))\mu(0)+\sum_{j=1}^{L}\lambda_R(\tilde{k}(k))\mu(|j-j_{\rm c}(k)|_{\rm circ})\Big)^2\Big] \quad (A.19)$$

$$= A_0 \exp\Big[\frac{1}{2}\sum_{k=\tilde{k}_s L+1}^{nL}\Big(\lambda_R(\tilde{k}(k))\Big)^2\times\Big(\mu(0)+\sum_{j=1}^{L}\mu(|j-j_{\rm c}(k)|_{\rm circ})\Big)^2\Big] \quad (A.20)$$

$$= A_0 \exp\Big[\frac{1}{2}\sum_{k=\tilde{k}_s L+1}^{nL}\Big(\lambda_R(\tilde{k}(k))\Big)^2\times\Big(\mu(0)+\sum_{\triangle j=0,\pm1,\ldots,\pm\frac{L-1}{2}}\mu(|\triangle j|)\Big)^2\Big] \quad (A.21)$$

$$= A_0 \exp\Big[\frac{1}{2}\sum_{k=\tilde{k}_s L+1}^{nL}\Big(\lambda_R(\tilde{k}(k))\Big)^2\times\Big(2\sum_{r=0,1,\ldots,\frac{L-1}{2}}\mu(r)\Big)^2\Big] \quad (A.22)$$

$$= A_0 \exp\Big[2L\sum_{\tilde{k}=\tilde{k}_s+1}^{n}\lambda_R^2(\tilde{k})\Big(\sum_{r=0}^{(L-1)/2}\mu(r)\Big)^2\Big]. \quad (A.23)$$

Since $\sum_{\tilde{k}=\tilde{k}_s+1}^{n}\lambda_R^2(\tilde{k}) \leq \sum_{\tilde{k}=\tilde{k}_s+1}^{\infty}\lambda_R^2(\tilde{k}) \leq P_0$, we have

$$A_n(\vec{\sigma}) \leq M_0 \coloneqq A_0 \exp\Big(2LQ(L)P_0\Big). \quad (A.24)$$

Therefore, we prove that, for any fixed $L > 0$, $A_n(\vec{\sigma})$ is upper bounded by M_0 for all $n > \tilde{k}_s$.

Using similar mathematical tricks with a modified range of k and $\tilde{k}$, for $n > \tilde{k}_s$ and $m \in \mathbb{N}$, we can upper bound $|A_{n+m}(\vec{\sigma})/A_n(\vec{\sigma})|$ by

$$\Big|\frac{A_{n+m}(\vec{\sigma})}{A_n(\vec{\sigma})}\Big| = \prod_{k=nL+1}^{(n+m)L}\Big|\cosh(\theta_k(\vec{\sigma}))\Big| \quad (A.25)$$

$$\leq \exp\Big[2LQ(L)\sum_{\tilde{k}=n+1}^{n+m}\lambda_R^2(\tilde{k})\Big]. \quad (A.26)$$

We can also lower bound $|A_{n+m}(\vec{\sigma})/A_n(\vec{\sigma})|$ by

$$\Big|\frac{A_{n+m}(\vec{\sigma})}{A_n(\vec{\sigma})}\Big| = \prod_{k=nL+1}^{(n+m)L} \frac{1}{2}\sqrt{e^{2u_k(\vec{\sigma})} + e^{-2u_k(\vec{\sigma})} + 2\cos(2v_k(\vec{\sigma}))} \tag{A.27}$$

$$\geq \prod_{k=nL+1}^{(n+m)L} \frac{1}{2}\sqrt{2 + 2\cos(2v_k(\vec{\sigma}))} \tag{A.28}$$

$$= \prod_{k=nL+1}^{(n+m)L} \Big|\cos(|v_k(\vec{\sigma})|)\Big|. \tag{A.29}$$

Since

$$V_k = |\operatorname{Im}(b_k^{(L)})| + \sum_{j=1}^{L} |\operatorname{Im}(W_{j,k}^{(L)})| \tag{A.30}$$

$$\leq 2\lambda_I(\tilde{k}(k)) \sum_{r=0}^{(L-1)/2} \mu(r), \tag{A.31}$$

and $\lambda_I(\tilde{k}(k)) \to 0$ as $k \to \infty$, we know that $V_k \to 0$ as $k \to \infty$. Define $n_I = n_I(L) = \min\{n \geq \tilde{k}_s, n \in \mathbb{N} : \forall k \geq n_I L + 1, V_k \leq \pi/3\}$. Note that the notation n_I defined here will be used throughout this paper. Using the fact that $|\cos(x)| \geq \exp(-\beta_1^2 x^2/2)$ with $\beta_1 = 3\sqrt{2\ln 2}/\pi$ for all $x \in [0, \pi/3]$, for $n \geq n_I$,

$$\Big|\frac{A_{n+m}(\vec{\sigma})}{A_n(\vec{\sigma})}\Big| \geq \prod_{k=nL+1}^{(n+m)L} |\cos(V_k)| \tag{A.32}$$

$$\geq \exp(-\frac{1}{2}\beta_1^2 \sum_{k=nL+1}^{(n+m)L} V_k^2) \tag{A.33}$$

$$= \exp\Big(-\frac{1}{2}\beta_1^2 \sum_{k=nL+1}^{(n+m)L} (|\operatorname{Im}(b_k^{(L)})| + \sum_{j=1}^{L} |\operatorname{Im}(W_{j,k}^{(L)})|)^2\Big) \tag{A.34}$$

$$\geq \exp\Big[-2\beta_1^2 LQ(L) \sum_{\tilde{k}=n+1}^{n+m} \lambda_I^2(\tilde{k})\Big]. \tag{A.35}$$

Since $P(\tilde{k}_s)$ converges to a constant $P_0/(1+\beta_1^2)$ and all terms are nonnegative, we have $P(m') \to 0$ as $m' \to \infty$. So $\sum_{\tilde{k}=m'+1}^{\infty} \lambda_R^2(\tilde{k})$ and $\sum_{\tilde{k}=m'+1}^{\infty} \lambda_I^2(\tilde{k})$ also approach 0 as $m' \to \infty$.

Therefore, for any sufficiently small $\varepsilon > 0$, there exists $n_1 \geq n_I \geq \tilde{k}_s$ such that, for all $m' > n_1$,

$$\sum_{\tilde{k}=m'+1}^{\infty} \lambda_R^2(\tilde{k}) < \ln(1+\varepsilon/M_0)/(2LQ(L)), \tag{A.36}$$

$$\sum_{\tilde{k}=m'+1}^{\infty} \lambda_I^2(\tilde{k}) < \ln(1/(1-\varepsilon/M_0))/(2\beta_1^2 LQ(L)). \tag{A.37}$$

Then, for all $n > n_1$ and $m \in \mathbb{N}$,

$$\Big|\frac{A_{n+m}(\vec{\sigma})}{A_n(\vec{\sigma})}\Big| \leq \exp\Big[2LQ(L)\sum_{\tilde{k}=n+1}^{\infty} \lambda_R^2(\tilde{k})\Big] \tag{A.38}$$

$$< 1+\varepsilon/M_0, \tag{A.39}$$

$$\Big|\frac{A_{n+m}(\vec{\sigma})}{A_n(\vec{\sigma})}\Big| \geq \exp\Big[-2\beta_1^2 LQ(L)\sum_{\tilde{k}=n+1}^{\infty} \lambda_I^2(\tilde{k})\Big] \tag{A.40}$$

$$> 1-\varepsilon/M_0, \tag{A.41}$$

which implies $|A_{n+m}(\vec{\sigma})/A_n(\vec{\sigma}) - 1| < \varepsilon/M_0$.

Therefore,

$$|A_{n+m}(\vec{\sigma}) - A_n(\vec{\sigma})| = A_n(\vec{\sigma})\Big|\frac{A_{n+m}(\vec{\sigma})}{A_n(\vec{\sigma})} - 1\Big| \tag{A.42}$$

$$< M_0\frac{\varepsilon}{M_0} = \varepsilon. \tag{A.43}$$

Thus, we prove that the sequence $\{A_n(\vec{\sigma}) : n \in \mathbb{N}\}$ is a Cauchy sequence in $\mathbb{R}$. So $A_n(\vec{\sigma})$

must converge to some nonnegative constant $A_\infty(\vec{\sigma}) \in \mathbb{R}$ as $n \to \infty$.

For $n > \tilde{k}_s$ and $m \in \mathbb{N}$, we have

$$
\begin{aligned}
&|\phi_{n+m}(\vec{\sigma}) - \phi_n(\vec{\sigma})| \\
&= \Big| \arg\Big(\frac{\psi^{(L,(n+m)L)}(\vec{\sigma})}{\psi^{(L,nL)}(\vec{\sigma})} \Big) \Big| && \text{(A.44)} \\
&= \Big| \arg\Big(\prod_{k=nL+1}^{(n+m)L} \cosh(u_k(\vec{\sigma})) \cos v_k(\vec{\sigma}) + i \sinh(u_k(\vec{\sigma})) \sin v_k(\vec{\sigma}) \Big) \Big| && \text{(A.45)} \\
&= \Big| \sum_{k=nL+1}^{(n+m)L} \arg\Big(\cosh(u_k(\vec{\sigma})) \cos v_k(\vec{\sigma}) + i \sinh(u_k(\vec{\sigma})) \sin v_k(\vec{\sigma}) \Big) \Big|. && \text{(A.46)}
\end{aligned}
$$

For $n \geq n_I > \tilde{k}_s$, $|v_k(\vec{\sigma})| \leq V_k \leq \pi/3$, so $\cosh(u_k(\vec{\sigma})) \cos v_k(\vec{\sigma}) > 0$, which implies that the argument of each term is in the range $(-\pi/2, \pi/2)$. So we have

$$
\begin{aligned}
&|\phi_{n+m}(\vec{\sigma}) - \phi_n(\vec{\sigma})| \\
&= \Big| \sum_{k=nL+1}^{(n+m)L} \arctan\Big(\tanh(u_k(\vec{\sigma})) \tan(v_k(\vec{\sigma})) \Big) \Big| && \text{(A.47)} \\
&\leq \sum_{k=nL+1}^{(n+m)L} \Big| \arctan\Big(\tanh(u_k(\vec{\sigma})) \tan(v_k(\vec{\sigma})) \Big) \Big| && \text{(A.48)} \\
&= \sum_{k=nL+1}^{(n+m)L} \arctan\Big(\tanh(|u_k(\vec{\sigma})|) \tan(|v_k(\vec{\sigma})|) \Big). && \text{(A.49)}
\end{aligned}
$$

Using

$$
\arctan(\tanh(x) \tan(y)) \leq \beta_2 \tanh(x) y \leq \beta_2 x y \qquad \text{(A.50)}
$$

with $\beta_2 = 3\sqrt{3}/\pi$ for all $x \geq 0$ and $0 \leq y \leq \pi/3$, we can get

$$
\begin{aligned}
&|\phi_{n+m}(\vec{\sigma}) - \phi_n(\vec{\sigma})| \\
&\leq \beta_2 \sum_{k=nL+1}^{(n+m)L} |u_k(\vec{\sigma})||v_k(\vec{\sigma})| && \text{(A.51)} \\
&\leq \beta_2 \sum_{k=nL+1}^{(n+m)L} U_k V_k && \text{(A.52)} \\
&= \beta_2 \sum_{k=nL+1}^{(n+m)L} \Big(|\operatorname{Re}(b_k^{(L)})| + \sum_{j=1}^{L} |\operatorname{Re}(W_{j,k}^{(L)})| \Big) \times \Big(|\operatorname{Im}(b_k^{(L)})| + \sum_{j=1}^{L} |\operatorname{Im}(W_{j,k}^{(L)})| \Big) && \text{(A.53)} \\
&\leq \beta_2 \Big[4L \sum_{\tilde{k}=n+1}^{n+m} \lambda_R(\tilde{k}) \lambda_I(\tilde{k}) \Big(\sum_{r=0}^{(L-1)/2} \mu(r) \Big)^2 \Big] && \text{(A.54)} \\
&= 4\beta_2 LQ(L) \sum_{\tilde{k}=n+1}^{n+m} \lambda_R(\tilde{k}) \lambda_I(\tilde{k}). && \text{(A.55)}
\end{aligned}
$$

Again, since $\sum_{\tilde{k}=m'+1}^{\infty} \lambda_R^2(\tilde{k})$ and $\sum_{\tilde{k}=m'+1}^{\infty} \lambda_I^2(\tilde{k})$ approach 0 as $m' \to \infty$,

$$
\sum_{\tilde{k}=m'+1}^{\infty} \lambda_R(\tilde{k}) \lambda_I(\tilde{k}) \leq \frac{1}{2} \sum_{\tilde{k}=m'+1}^{\infty} (\lambda_R^2(\tilde{k}) + \lambda_I^2(\tilde{k})) \tag{A.56}
$$

also approaches 0 as $m' \to \infty$.

Therefore, for any sufficiently small $\varepsilon > 0$, there exists $n_1 \geq n_I \geq \tilde{k}_s$ such that, for all $m' > n_1$,

$$
\sum_{\tilde{k}=m'+1}^{\infty} \lambda_R(\tilde{k}) \lambda_I(\tilde{k}) < \varepsilon/(4\beta_2 LQ(L)). \tag{A.57}
$$

Then, for all $n > n_1$ and $m \in \mathbb{N}$,

$$
|\phi_{n+m}(\vec{\sigma}) - \phi_n(\vec{\sigma})| < 4\beta_2 LQ(L) \frac{\varepsilon}{4\beta_2 LQ(L)} = \varepsilon. \tag{A.58}
$$

Thus, we prove that the sequence $\{\phi_n(\vec{\sigma}) : n \in \mathbb{N}\}$ is a Cauchy sequence in $\mathbb{R}$. So $\phi_n(\vec{\sigma})$ must converge to some constant $\phi_\infty(\vec{\sigma}) \in \mathbb{R}$ as $n \to \infty$.

Combining the above two conclusions, we can get that, for any sufficiently small $\varepsilon > 0$, there exists $n_3 \geq n_I \geq \tilde{k}_s$ such that, for all $n > n_3$ and $m \in \mathbb{N}$, $|A_{n+m}(\vec{\sigma}) - A_n(\vec{\sigma})| < \varepsilon/\sqrt{2}$ and $|\phi_{n+m}(\vec{\sigma}) - \phi_n(\vec{\sigma})| < \arccos(1 - \frac{\varepsilon^2}{4M_0^2})$. So

$$
\begin{aligned}
&\left|\psi^{(L,(n+m)L)}(\vec{\sigma}) - \psi^{(L,nL)}(\vec{\sigma})\right| \\
&= \left[(A_{n+m}(\vec{\sigma}) - A_n(\vec{\sigma}))^2 + 2A_{n+m}(\vec{\sigma})A_n(\vec{\sigma})\Big(1 - \cos(\phi_{n+m}(\vec{\sigma}) - \phi_n(\vec{\sigma}))\Big)\right]^{1/2} \quad \text{(A.59)} \\
&< \left[\frac{1}{2}\varepsilon^2 + 2M_0^2\frac{\varepsilon^2}{4M_0^2}\right]^{1/2} = \varepsilon. \quad \text{(A.60)}
\end{aligned}
$$

Thus, we prove that the sequence $\{\psi^{(L,nL)}(\vec{\sigma}) : n \in \mathbb{N}\}$ is a Cauchy sequence in the metric space $(\mathbb{C}, d)$, and $\psi^{(L,nL)}(\vec{\sigma})$ actually converges to the constant $A_\infty(\vec{\sigma})e^{i\phi_\infty(\vec{\sigma})} \in \mathbb{C}$ as $n \to \infty$.

□

A.2 Proof of upper bounds on truncation errors for LRFD RBMs

In this section, we provide a proof of the upper bounds on the two types of truncation errors for LRFD RBMs given in Sec. 2.2.2.

We give the proof for the first-type truncation errors as follows.

Proof.

$$
\begin{aligned}
&\| |\tilde{\Psi}^{(L,\infty)}\rangle - |\tilde{\Psi}^{(L,N_h)}\rangle \|^2 \\
&= 2 - (\langle \tilde{\Psi}^{(L,\infty)} | \tilde{\Psi}^{(L,N_h)} \rangle + \text{C.c.}) && \text{(A.61)} \\
&= 2 - \frac{(\langle \Psi^{(L,\infty)} | \Psi^{(L,N_h)} \rangle + \text{C.c.})}{\sqrt{\langle \Psi^{(L,\infty)} | \Psi^{(L,\infty)} \rangle \langle \Psi^{(L,N_h)} | \Psi^{(L,N_h)} \rangle}} && \text{(A.62)} \\
&= 2 - \frac{\sum_{\vec{\sigma}} (\psi^{(L,\infty)}(\vec{\sigma})^* \psi^{(L,N_h)}(\vec{\sigma}) + \text{C.c.})}{\sqrt{\sum_{\vec{\sigma}} |\psi^{(L,\infty)}(\vec{\sigma})|^2 \sum_{\vec{\sigma}} |\psi^{(L,N_h)}(\vec{\sigma})|^2}} && \text{(A.63)} \\
&= 2 - \frac{2 \sum_{\vec{\sigma}} |\psi^{(L,N_h)}(\vec{\sigma})|^2 \, \text{Re} \left(\frac{\psi^{(L,\infty)}(\vec{\sigma})}{\psi^{(L,N_h)}(\vec{\sigma})} \right)}{\sqrt{\sum_{\vec{\sigma}} |\psi^{(L,\infty)}(\vec{\sigma})|^2 \sum_{\vec{\sigma}} |\psi^{(L,N_h)}(\vec{\sigma})|^2}}. && \text{(A.64)}
\end{aligned}
$$

Define

$$
\Theta = \Theta(L, N_h) = \max_{\vec{\sigma}} \{ \Big| \arg \left(\frac{\psi^{(L,\infty)}(\vec{\sigma})}{\psi^{(L,N_h)}(\vec{\sigma})} \right) \Big| \}. \tag{A.65}
$$

For $N_h/L \geq n_I(L) > \tilde{k}_s$,

$$
\Big| \arg \left(\frac{\psi^{(L,\infty)}(\vec{\sigma})}{\psi^{(L,N_h)}(\vec{\sigma})} \right) \Big| \leq 4\beta_2 L Q(L) \sum_{\tilde{k}=N_h/L+1}^{\infty} \lambda_R(\tilde{k}) \lambda_I(\tilde{k}). \tag{A.66}
$$

Therefore,

$$
\Theta \leq 4\beta_2 L Q(L) \sum_{\tilde{k}=N_h/L+1}^{\infty} \lambda_R(\tilde{k}) \lambda_I(\tilde{k}). \tag{A.67}
$$

Define

$$R_1 = R_1(L, N_h) = \max_{\vec{\sigma}}\{\Big|\frac{\psi^{(L,\infty)}(\vec{\sigma})}{\psi^{(L,N_h)}(\vec{\sigma})}\Big|^2\}, \tag{A.68}$$

$$R_2 = R_2(L, N_h) = \min_{\vec{\sigma}}\{\Big|\frac{\psi^{(L,\infty)}(\vec{\sigma})}{\psi^{(L,N_h)}(\vec{\sigma})}\Big|^2\}. \tag{A.69}$$

We have

$$R_1 \leq \exp\Big[4LQ(L)\sum_{\tilde{k}=N_h/L+1}^{\infty}\lambda_R^2(\tilde{k})\Big], \tag{A.70}$$

$$R_2 \geq \exp\Big[-4\beta_1^2 LQ(L)\sum_{\tilde{k}=N_h/L+1}^{\infty}\lambda_I^2(\tilde{k})\Big]. \tag{A.71}$$

Consider

$$\mathrm{Re}\Big(\frac{\psi^{(L,\infty)}(\vec{\sigma})}{\psi^{(L,N_h)}(\vec{\sigma})}\Big) \geq \sqrt{R_2}\cos(\Theta), \tag{A.72}$$

$$\sum_{\vec{\sigma}}|\psi^{(L,\infty)}(\vec{\sigma})|^2 \leq R_1\sum_{\vec{\sigma}}|\psi^{(L,N_h)}(\vec{\sigma})|^2. \tag{A.73}$$

We can get

$$
\begin{aligned}
&\| |\tilde{\Psi}^{(L,\infty)}\rangle - |\tilde{\Psi}^{(L,N_h)}\rangle \|^2 \\
&\leq 2 - \frac{2\sum_{\vec{\sigma}} |\psi^{(L,N_h)}(\vec{\sigma})|^2 \sqrt{R_2}\cos(\Theta)}{\sqrt{R_1\Big(\sum_{\vec{\sigma}} |\psi^{(L,N_h)}(\vec{\sigma})|^2\Big)^2}}. && \text{(A.74)} \\
&= 2 - 2\sqrt{\frac{R_2}{R_1}}\cos(\Theta) && \text{(A.75)} \\
&\leq 2 - 2\exp\Big[-2LQ(L)\sum_{\tilde{k}=N_h/L+1}^{\infty}\Big(\lambda_R^2(\tilde{k}) + \beta_1^2\lambda_I^2(\tilde{k})\Big)\Big] \\
&\quad \times \cos\Big[4\beta_2 LQ(L)\sum_{\tilde{k}=N_h/L+1}^{\infty}\lambda_R(\tilde{k})\lambda_I(\tilde{k})\Big]. && \text{(A.76)}
\end{aligned}
$$

For simplicity, we assume $\lambda_R(\tilde{k}) = \lambda_I(\tilde{k}) = \lambda(\tilde{k})$ which means that the real part and imaginary part of RBM parameters have the same decaying rate and define

$$
P(N_h/L) = \sum_{\tilde{k}=N_h/L+1}^{\infty}\lambda^2(\tilde{k}). \tag{A.77}
$$

Then

$$
\begin{aligned}
&\| |\tilde{\Psi}^{(L,\infty)}\rangle - |\tilde{\Psi}^{(L,N_h)}\rangle \|^2 \\
&\leq 2 - 2\exp\Big[-2(1+\beta_1^2)LQ(L)P(N_h/L)\Big]\cos\Big[4\beta_2 LQ(L)P(N_h/L)\Big] && \text{(A.78)} \\
&= F_1\Big(LQ(L)P(N_h/L)\Big), && \text{(A.79)}
\end{aligned}
$$

where

$$F_1(x) = 2 - 2\exp[-2(1+\beta_1^2)x]\cos(4\beta_2 x) \tag{A.80}$$

$$= 4(1+\beta_1^2)x + \mathcal{O}(x^2) \tag{A.81}$$

$$= \mathcal{O}(x) \quad (\text{as } x \to 0). \tag{A.82}$$

The idea of the proof is shown schematically in Fig. 2.2(c), where $\psi_{\text{full}}(\vec{\sigma})$ denotes the amplitude for the full LRFD RBM $\psi^{(L,\infty)}(\vec{\sigma})$ and $\psi_{\text{tr}}(\vec{\sigma})$ denotes the amplitude for the truncated LRFD RBM $\psi^{(L,N_h)}(\vec{\sigma})$. □

We give the proof for the second-type truncation errors as follows.

We have defined a Hermitian operator $\hat{B}$ of the form $\hat{B} = \bigotimes_{j=1}^{L} \hat{\sigma}_j^{(m_j)}$ where $\bigotimes$ is the tensor product symbol, $m_j \in \{0,1,2,3\}$, $\hat{\sigma}_j^{(0)} = I_{2\times 2}$ is the identity matrix, and $\{\hat{\sigma}_j^{(1)}, \hat{\sigma}_j^{(2)}, \hat{\sigma}_j^{(3)}\}$ denote the Pauli matrices.

Proof.

$$\begin{aligned}
&|\langle \hat{B} \rangle^{(L,\infty)} - \langle \hat{B} \rangle^{(L,N_h)}| \\
&= \Big| \frac{\langle \Psi^{(L,\infty)} | \hat{B} | \Psi^{(L,\infty)} \rangle}{\langle \Psi^{(L,\infty)} | \Psi^{(L,\infty)} \rangle} - \frac{\langle \Psi^{(L,N_h)} | \hat{B} | \Psi^{(L,N_h)} \rangle}{\langle \Psi^{(L,N_h)} | \Psi^{(L,N_h)} \rangle} \Big| && \text{(A.83)} \\
&\le \Big| \frac{\langle \Psi^{(L,\infty)} | \hat{B} | \Psi^{(L,\infty)} \rangle}{\langle \Psi^{(L,\infty)} | \Psi^{(L,\infty)} \rangle} - \frac{\langle \Psi^{(L,\infty)} | \hat{B} | \Psi^{(L,\infty)} \rangle}{\langle \Psi^{(L,N_h)} | \Psi^{(L,N_h)} \rangle} \Big| \\
&\quad + \Big| \frac{\langle \Psi^{(L,\infty)} | \hat{B} | \Psi^{(L,\infty)} \rangle}{\langle \Psi^{(L,N_h)} | \Psi^{(L,N_h)} \rangle} - \frac{\langle \Psi^{(L,N_h)} | \hat{B} | \Psi^{(L,N_h)} \rangle}{\langle \Psi^{(L,N_h)} | \Psi^{(L,N_h)} \rangle} \Big| && \text{(A.84)} \\
&= G_1 + G_2, && \text{(A.85)}
\end{aligned}$$

where

$$G_1 = \Big| \frac{\langle \Psi^{(L,\infty)} | \Psi^{(L,\infty)} \rangle}{\langle \Psi^{(L,N_h)} | \Psi^{(L,N_h)} \rangle} - 1 \Big| \Big| \langle \hat{B} \rangle^{(L,\infty)} \Big| \tag{A.86}$$

$$G_2 = \Big| \langle \Psi^{(L,\infty)} | \hat{B} | \Psi^{(L,\infty)} \rangle - \langle \Psi^{(L,N_h)} | \hat{B} | \Psi^{(L,N_h)} \rangle \Big| (\langle \Psi^{(L,N_h)} | \Psi^{(L,N_h)} \rangle)^{-1} \tag{A.87}$$

$$\begin{aligned} &= \sum_{\vec{\sigma}_1} \sum_{\vec{\sigma}_2} \Big| B_{\vec{\sigma}_1 \vec{\sigma}_2} \psi^{(L,N_h)}(\vec{\sigma}_1)^* \psi^{(L,N_h)}(\vec{\sigma}_2) \Big(\frac{\psi^{(L,\infty)}(\vec{\sigma}_1)^* \psi^{(L,\infty)}(\vec{\sigma}_2)}{\psi^{(L,N_h)}(\vec{\sigma}_1)^* \psi^{(L,N_h)}(\vec{\sigma}_2)} - 1 \Big) \Big| \\ &\quad \times \Big(\sum_{\vec{\sigma}} |\psi^{(L,N_h)}(\vec{\sigma})|^2 \Big)^{-1}. \end{aligned} \tag{A.88}$$

Considering $R_2 \leq \sum_{\vec{\sigma}} |\psi^{(L,\infty)}(\vec{\sigma})|^2 / \sum_{\vec{\sigma}} |\psi^{(L,N_h)}(\vec{\sigma})|^2 \leq R_1$ and $|\langle \hat{B} \rangle^{(L,\infty)}| \leq 1$, we can get

$$G_1 \leq \max\{|R_1 - 1|, |1 - R_2|\} \tag{A.89}$$

Define

$$\xi = \max_{(\vec{\sigma}_1, \vec{\sigma}_2)} \Big| \frac{\psi^{(L,\infty)}(\vec{\sigma}_1)^* \psi^{(L,\infty)}(\vec{\sigma}_2)}{\psi^{(L,N_h)}(\vec{\sigma}_1)^* \psi^{(L,N_h)}(\vec{\sigma}_2)} - 1 \Big|. \tag{A.90}$$

Then

$$G_2 \leq \xi \sum_{\vec{\sigma}_1} \sum_{\vec{\sigma}_2} \Big| B_{\vec{\sigma}_1 \vec{\sigma}_2} \Big| \Big| \psi^{(L,N_h)}(\vec{\sigma}_1)^* \Big| \Big| \psi^{(L,N_h)}(\vec{\sigma}_2) \Big| \Big(\sum_{\vec{\sigma}} |\psi^{(L,N_h)}(\vec{\sigma})|^2 \Big)^{-1} \tag{A.91}$$

$$= \xi \sum_{\vec{\sigma}_2} \Big| \psi^{(L,N_h)}(\hat{B}(\vec{\sigma}_2))^* \Big| \Big| \psi^{(L,N_h)}(\vec{\sigma}_2) \Big| \Big(\sum_{\vec{\sigma}} |\psi^{(L,N_h)}(\vec{\sigma})|^2 \Big)^{-1}. \tag{A.92}$$

Using the Cauchy-Schwarz inequality,

$$\begin{aligned}
&\sum_{\vec{\sigma}_2} \Big|\psi^{(L,N_h)}(\hat{B}(\vec{\sigma}_2))^*\Big|\Big|\psi^{(L,N_h)}(\vec{\sigma}_2)\Big| \\
&\leq \Big(\sum_{\vec{\sigma}_2} \Big|\psi^{(L,N_h)}(\hat{B}(\vec{\sigma}_2))^*\Big|^2 \sum_{\vec{\sigma}_2} \Big|\psi^{(L,N_h)}(\vec{\sigma}_2)\Big|^2\Big)^{1/2} &\text{(A.93)}\\
&= \sum_{\vec{\sigma}} |\psi^{(L,N_h)}(\vec{\sigma})|^2. &\text{(A.94)}
\end{aligned}$$

Therefore, combined with the geometric features, we can get

$$\begin{aligned}
G_2 &\leq \xi &\text{(A.95)}\\
&\leq \max\Big\{\sqrt{(R_1\cos(2\Theta)-1)^2+R_1^2\sin^2(2\Theta)}, \\
&\quad \sqrt{(R_2\cos(2\Theta)-1)^2+R_2^2\sin^2(2\Theta)}, |R_1-1|, |1-R_2|\Big\} &\text{(A.96)}\\
&= \max\Big\{\sqrt{R_1^2-2R_1\cos(2\Theta)+1}, \sqrt{R_2^2-2R_2\cos(2\Theta)+1}\Big\}. &\text{(A.97)}
\end{aligned}$$

Therefore,

$$\begin{aligned}
&|\langle\hat{B}\rangle^{(L,\infty)} - \langle\hat{B}\rangle^{(L,N_h)}| \\
&\leq \max\{|R_1-1|, |1-R_2|\} \\
&\quad + \max\Big\{\sqrt{R_1^2-2R_1\cos(2\Theta)+1}, \sqrt{R_2^2-2R_2\cos(2\Theta)+1}\Big\} &\text{(A.98)}\\
&\leq F_2\Big(LQ(L)P(N_h/L)\Big), &\text{(A.99)}
\end{aligned}$$

where

$$F_2(x) = \max\{|\exp(4x) - 1|, |1 - \exp(-4\beta_1^2 x)|\}$$
$$+ \max\{\Big[\exp(8x) - 2\exp(4x)\cos(8\beta_2 x) + 1\Big]^{1/2},$$
$$\Big[\exp(-8\beta_1^2 x) - 2\exp(-4\beta_1^2 x)\cos(8\beta_2 x) + 1\Big]^{1/2}\} \tag{A.100}$$
$$= [4\beta_1^2 + 4(\beta_1^4 + 4\beta_2^2)^{1/2}]x + \mathcal{O}(x^{3/2}) \tag{A.101}$$
$$= \mathcal{O}(x) \quad (\text{as } x \to 0). \tag{A.102}$$

□

A.3 Proof of spin correlation formula

In this section, we provide a proof of the spin-correlation formulas in Sec. 2.2.4.

Proof. We assume that all RBM parameters a_j's, b_k's and $W_{j,k}$'s are no larger than ε_1, and $\varepsilon_1 \ll 1/L$, $\varepsilon_1 \ll 1/N_h$. In this section, we will not explicitly write the superscript "$^{(L,N_h)}$" and assume that the RBM just has a finite number (N_h) of hidden nodes. Then we can get the Taylor series expansion of $|\psi(\vec{\sigma})|^2$ in a small-parameter regime:

$$|\psi(\vec{\sigma})|^2 = \exp(\sum_{j=1}^{L} 2\,\mathrm{Re}(a_j)\sigma_j) \prod_{k=1}^{N_h} |\cosh(\theta_k)|^2 \tag{A.103}$$
$$= F^{(0)} + \sum_{j_1 \in \{1,2,\ldots,L\}} F_{j_1}^{(1)} \sigma_{j_1} + \ldots \tag{A.104}$$
$$= \sum_{n=0}^{L} \sum_{1 \le j_1 < \cdots < j_n \le L} F_{j_1,\ldots,j_n}^{(n)} \prod_{j=j_1}^{j_n} \sigma_j. \tag{A.105}$$

Then, the spin correlation between spin 1 and spin $1 + r$ in the longitudinal (z) direction

$C^z_{\text{unnorm}}(r)$ is

$$C^z_{\text{unnorm}}(r) = \langle\Psi|\hat{\sigma}^z_1\hat{\sigma}^z_{1+r}|\Psi\rangle \tag{A.106}$$

$$= \sum_{\sigma_1=\pm1}\cdots\sum_{\sigma_L=\pm1}\sigma_1\sigma_{1+r}|\psi(\vec{\sigma})|^2 \tag{A.107}$$

$$= \sum_{\sigma_1=\pm1}\sum_{\sigma_{1+r}=\pm1}\sigma_1\sigma_{1+r}\sum_{\vec{\sigma}\setminus\{\sigma_1,\sigma_{1+r}\}}\sum_{n=0}^{L}\sum_{1\le j_1<\cdots<j_n\le L}F^{(n)}_{j_1,\ldots,j_n}\prod_{j=j_1}^{j_n}\sigma_j \tag{A.108}$$

$$\begin{aligned}= \sum_{\sigma_1=\pm1}\sum_{\sigma_{1+r}=\pm1}\sigma_1\sigma_{1+r}\sum_{\vec{\sigma}\setminus\{\sigma_1,\sigma_{1+r}\}}&(F^{(0)}\\ &+\sum_{1\le j_1\le L}F^{(1)}_{j_1}\sigma_{j_1}+\sum_{1\le j_1<j_2\le L}F^{(n)}_{j_1,j_2}\sigma_{j_1}\sigma_{j_2}+\mathcal{O}(\varepsilon_1^3)).\end{aligned} \tag{A.109}$$

By applying Taylor series expansions to each term, we can get

$$\begin{aligned}&\exp(\sum_{j=1}^{L}2\,\mathrm{Re}(a_j)\sigma_j)\\ &=\left(1+2\sum_{j=1}^{L}\mathrm{Re}(a_j)^2\right)+2\sum_{j=1}^{L}\mathrm{Re}(a_j)\sigma_j\\ &\quad+2\sum_{j_1=1}^{L}\sum_{1\le j_2\le L,j_2\ne j_1}\mathrm{Re}(a_{j_1})\,\mathrm{Re}(a_{j_2})\sigma_{j_1}\sigma_{j_2}+\mathcal{O}(\varepsilon_1^3)\quad(\text{as }\varepsilon_1\to0).\end{aligned} \tag{A.110}$$

$$\theta_k^2 = (b_k+\sum_{j=1}^{L}W_{j,k}\sigma_j)^2 \tag{A.111}$$

$$= (b_k^2+\sum_{j=1}^{L}W_{j,k}^2)+2b_k\sum_{j=1}^{L}W_{j,k}\sigma_j+\sum_{j_1=1}^{L}\sum_{1\le j_2\le L,j_2\ne j_1}W_{j_1,k}W_{j_2,k}\sigma_{j_1}\sigma_{j_2}. \tag{A.112}$$

$$|\cosh(\theta_k)|^2 = 1 + \frac{1}{2}\theta_k^2 + \frac{1}{2}(\theta_k^*)^2 + \mathcal{O}(\varepsilon_1^4) \tag{A.113}$$

$$= 1 + \mathrm{Re}\Big((b_k^2 + \sum_{j=1}^{L} W_{j,k}^2) + 2b_k \sum_{j=1}^{L} W_{j,k}\sigma_j + \sum_{j_1=1}^{L} \sum_{1\le j_2\le L, j_2\neq j_1} W_{j_1,k}W_{j_2,k}\sigma_{j_1}\sigma_{j_2}\Big) + \mathcal{O}(\varepsilon_1^4) \tag{A.114}$$

$$= \Big[1 + \mathrm{Re}(b_k^2 + \sum_{j=1}^{L} W_{j,k}^2)\Big] + 2\,\mathrm{Re}(b_k \sum_{j=1}^{L} W_{j,k})\sigma_j + \mathrm{Re}(\sum_{j_1=1}^{L} \sum_{1\le j_2\le L, j_2\neq j_1} W_{j_1,k}W_{j_2,k})\sigma_{j_1}\sigma_{j_2} + \mathcal{O}(\varepsilon_1^4) \quad (\text{as } \varepsilon_1 \to 0). \tag{A.115}$$

Therefore,

$$F_{1,1+r}^{(1)} = 2\sum_{k=1}^{N_h} \mathrm{Re}(W_{1,k}W_{1+r,k}) + 4\,\mathrm{Re}(a_1)\,\mathrm{Re}(a_{1+r}) + \mathcal{O}(\varepsilon_1^3) \quad (\text{as } \varepsilon_1 \to 0). \tag{A.116}$$

Therefore,

$$C_{\text{unnorm}}^{z}(r) = F_{1,1+r}^{(1)} \tag{A.117}$$

$$= 2\big(\mathrm{Re}(WW^T)\big)_{1,1+r} + 4\,\mathrm{Re}(a_1)\,\mathrm{Re}(a_{1+r}) + \mathcal{O}(\varepsilon_1^3) \quad (\text{as } \varepsilon_1 \to 0). \tag{A.118}$$

We construct an RBM by setting

$$W_{j,k} = \lambda(\tilde{k})\mu(|j - j_c|_{\text{circ}}), \tag{A.119}$$

$$a_j = 0, \quad b_k = 0 \tag{A.120}$$

for all $1 \leq j \leq L$ and $1 \leq k \leq N_h$. Then

$$C^z_{\text{unnorm}}(r) \approx 2(WW^T)_{1,1+r} \tag{A.121}$$

$$= \sum_{\tilde{k}=1}^{N_h/L} |\lambda(\tilde{k})|^2 \sum_{j_c=1}^{L} \mu(|1-j_c|_{\text{circ}})\mu(|1+r-j_c|_{\text{circ}}). \tag{A.122}$$

Considering the definition of $|\ldots|_{\text{circ}}$ and assume L is even without influencing the asymptotic analysis, the r-related part in $C^z_{\text{unnorm}}(r)$ is

$$G^z(r) = \sum_{j_c=1}^{L} \mu(|1-j_c|_{\text{circ}})\mu(|1+r-j_c|_{\text{circ}}) \tag{A.123}$$

$$= \sum_{j_c=1}^{r-1} \mu(j_c)\mu(r-j_c) + \sum_{j_c=L/2+1-r}^{L/2-1} \mu(j_c)\mu(L-r-j_c)$$

$$+2 \sum_{j_c=r+1}^{L/2} \mu(j_c)\mu(j_c-r) + 2\mu(0)\mu(r). \tag{A.124}$$

Considering the variation of each term with varying j_c in the above summation in a ring-shaped geometry corresponding to the periodic boundary conditions, we know that the contribution from the first two summations dominates $G^z(r)$ as $r \to L/2$, and the contribution from the second summation is no larger than that from the first one. So we can just focus on the contribution from the first summation when analyzing the asymptotic long-range behavior of $G^z(r)$ as $L \to \infty$ and $r \to L/2$.

For $\mu(r) = \delta_Q / r$,

$$\sum_{j_c=1}^{r-1} \mu(j_c)\mu(r-j_c) = \delta_Q^2 \sum_{j_c=1}^{r-1} \frac{1}{j_c}\frac{1}{r-j_c} = \delta_Q^2 \sum_{j_c=1}^{r-1} \frac{1}{r}(\frac{1}{j_c} + \frac{1}{r-j_c}) = \frac{2\delta_Q^2}{r} \sum_{j_c=1}^{r-1} \frac{1}{j_c} \quad \text{(A.125)}$$

$$= \frac{2\delta_Q^2}{r}(\ln(r-1) - \gamma + \mathcal{O}(\frac{1}{r})) \quad \text{(A.126)}$$

$$= 2\delta_Q^2 \frac{\ln r}{r} + \mathcal{O}(\frac{1}{r}) \quad (\text{as } L \to \infty, r \to L/2). \quad \text{(A.127)}$$

Therefore,

$$G^z(r) \geq 2\delta_Q^2 \frac{\ln r}{r} + \mathcal{O}(\frac{1}{r}) \quad (\text{as } L \to \infty, r \to L/2). \quad \text{(A.128)}$$

For $\mu(r) = \delta_Q / r^{\alpha_Q}$ with $\alpha_Q > 1$,

$$\sum_{j_c=1}^{r-1} \mu(j_c)\mu(r-j_c) = \delta_Q^2 \sum_{j_c=1}^{r-1} \frac{1}{j_c^{\alpha_Q}}\frac{1}{(r-j_c)^{\alpha_Q}} = \delta_Q^2 \sum_{j_c=1}^{r-1} \frac{1}{r^{\alpha_Q}}(\frac{1}{j_c} + \frac{1}{r-j_c})^{\alpha_Q} \quad \text{(A.129)}$$

$$\geq \frac{\delta_Q^2}{r^{\alpha_Q}} \sum_{j_c=1}^{r-1} (\frac{1}{j_c^{\alpha_Q}} + \frac{1}{(r-j_c)^{\alpha_Q}}) \quad \text{(A.130)}$$

$$= \frac{2\delta_Q^2}{r^{\alpha_Q}} \sum_{j_c=1}^{r-1} \frac{1}{j_c^{\alpha_Q}} = \frac{2\delta_Q^2}{r^{\alpha_Q}}(\zeta(\alpha_Q) - \mathcal{O}(\frac{1}{(r-1)^{\alpha_Q}})) \quad \text{(A.131)}$$

$$= 2\delta_Q^2 \zeta(\alpha_Q) \frac{1}{r^{\alpha_Q}} + \mathcal{O}(\frac{1}{r^{2\alpha_Q}}) \quad (\text{as } L \to \infty, r \to L/2). \quad \text{(A.132)}$$

Therefore,

$$G^z(r) \geq 2\delta_Q^2 \zeta(\alpha_Q) \frac{1}{r^{\alpha_Q}} + \mathcal{O}(\frac{1}{r^{2\alpha_Q}}) \quad (\text{as } L \to \infty, r \to L/2). \quad \text{(A.133)}$$

For $\mu(r) = \delta_Q / r^{\alpha_Q}$ with $0 < \alpha_Q \leq \frac{1}{2}$,

$$G^z(\frac{L}{2}) = 2\sum_{j_c=1}^{\frac{L}{2}-1} \mu(j_c)\mu(\frac{L}{2} - j_c) + 2\mu(0)\mu(\frac{L}{2}) \geq 2\delta_Q^2 \sum_{j_c=1}^{L/2-1} \frac{1}{j_c^{\alpha_Q}} \frac{1}{(L/2 - j_c)^{\alpha_Q}} \quad \text{(A.134)}$$

$$\geq 2\delta_Q^2 \sum_{j_c=1}^{L/2-1} \frac{1}{[(j_c + L/2 - j_c)/2]^{2\alpha_Q}} = 2\delta_Q^2 \frac{L/2-1}{(L/4)^{2\alpha_Q}} \geq 2\delta_Q^2 \frac{L/2-1}{L/4} \quad \text{(A.135)}$$

$$= 4\delta_Q^2 + o(1) \quad (\text{as } L \to \infty, r \to L/2). \quad \text{(A.136)}$$

□

While the truncation error for a geometric series can be directly estimated, we derive an estimation of the order of the truncation error for a series with terms having a power-law-decaying form as follows. We know that [123], for any $s > 1$, $b \geq a \geq 1$, and $a, b \in \mathbb{N}$,

$$\sum_{n=a+1}^{b} \frac{1}{n^s} = \frac{a^{1-s} - b^{1-s}}{s-1} - s\int_a^b \frac{(x - \lfloor x \rfloor - \frac{1}{2})}{x^{s+1}} \mathrm{d}x + \frac{b^{-s} - a^{-s}}{2}. \quad \text{(A.137)}$$

Consider

$$|s\int_a^b \frac{(x - \lfloor x \rfloor - \frac{1}{2})}{x^{s+1}} \mathrm{d}x| \leq \frac{s}{2}\int_a^b \frac{1}{x^{s+1}} \mathrm{d}x = \frac{a^{-s} - b^{-s}}{2}. \quad \text{(A.138)}$$

Therefore,

$$\sum_{n=a+1}^{b} \frac{1}{n^s} = \frac{a^{1-s} - b^{1-s}}{s-1} + \mathcal{O}(a^{-s}) \quad (\text{as } a \to \infty). \quad \text{(A.139)}$$

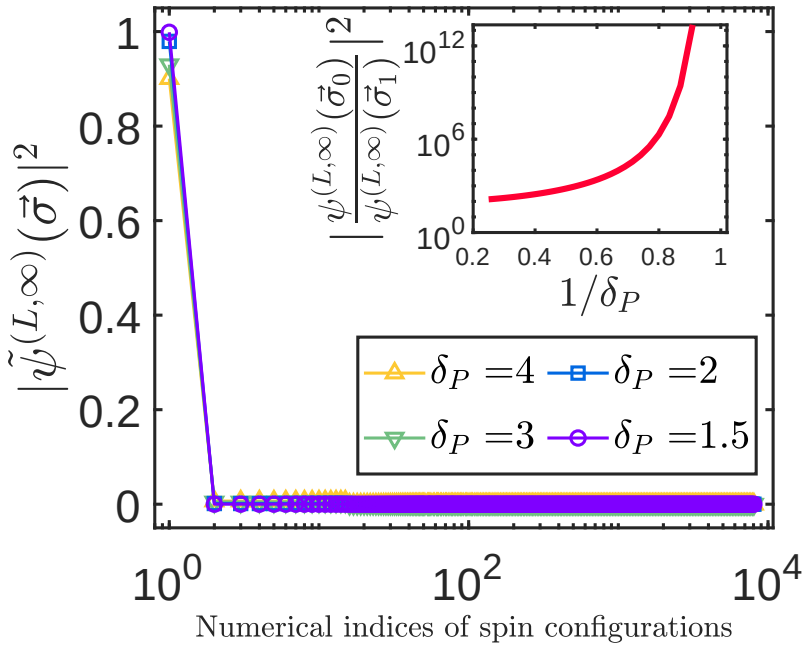


Figure A.1: Distribution of the square of normalized wave-function amplitudes in the spin-configuration space which can approximate the Kronecker delta function. The LRFD RBMs are constructed as Eqs. (A.141)–(A.143) show and $\mu_0 = 0.1$, $L = 13$ and $\lambda(\tilde{k}) = \delta_P^{-(\tilde{k}-1)}$ with $\delta_P > 1$. The horizontal axis denotes the numerical indices of all 2^L spin configurations which are sorted in a monotonically decreasing order of their corresponding amplitudes. The inset shows the ratio $|\psi^{(L,\infty)}(\vec{\sigma}_0)/\psi^{(L,\infty)}(\vec{\sigma}_1)|^2$ as a function of $1/\delta_P$.

The above results also hold for $b \to \infty$. Therefore,

$$\sum_{n=a+1}^{\infty} \frac{1}{n^s} = \frac{a^{1-s}}{s-1} + \mathcal{O}(a^{-s}) = \mathcal{O}(a^{1-s}) \quad (\text{as } a \to \infty). \tag{A.140}$$

A.4 LRFD RBMs approximating the Kronecker delta function

In this section, we construct LRFD RBMs which can approximate the quantum state with all spins pointing up in the z direction, which has a form of the Kronecker delta function, with arbitrary accuracy for any finite system size. The RBMs are constructed as

Eqs. (2.24)–(2.26) show with

$$\mu(r) = \mu_0 > 0 \quad (\text{for any } r \geq 0), \tag{A.141}$$

$$c_w = c_b = 1, \tag{A.142}$$

$$a_0 = 0. \tag{A.143}$$

Since all b_k and $W_{j,k}$ parameters for these RBMs are real positive numbers, the wave-function amplitude for spin configurations reaches its maximum at $\vec{\sigma}_0 = (1, 1, \ldots, 1)$ which corresponds to all spins up. We can get that, for any other spin configurations $\vec{\sigma}'$,

$$\begin{aligned} & |\psi^{(L,\infty)}(\vec{\sigma}_0)/\psi^{(L,\infty)}(\vec{\sigma}')| \\ \geq\ & |\psi^{(L,\infty)}(\vec{\sigma}_0)/\psi^{(L,\infty)}(\vec{\sigma}_1)| \qquad (A.144) \\ =\ & \prod_{\tilde{k}=1}^{\infty} \Big[\frac{\cosh((L+1)\mu_0\lambda(\tilde{k}))}{\cosh((L-1)\mu_0\lambda(\tilde{k}))}\Big]^L, \qquad (A.145) \end{aligned}$$

where $\vec{\sigma}_1 = (1, 1, \ldots, 1, -1)$ refers to the configuration obtained by flipping the last spin in $\vec{\sigma}_0$. Let $x = (L-1)\mu_0\lambda(\tilde{k})$ and $\triangle x = 2\mu_0\lambda(\tilde{k})$. By performing the Taylor series expansion of $\cosh(x + \triangle x)$ about x and comparing the leading-order terms, we can get that there exist constants $0 < \beta_3 < 1$ and $x_0 > 0$ such that, for any $0 < \triangle x < x < x_0$,

$$\frac{\cosh(x + \triangle x)}{\cosh(x)} \geq 1 + \tanh(x)\triangle x \geq e^{\beta_3 x \triangle x}. \tag{A.146}$$

For any fixed L and μ_0, there exists $\tilde{k}_0 \in \mathbb{N}$ such that $(L-1)\mu_0\lambda(\tilde{k}) < x_0$ for all $\tilde{k} > \tilde{k}_0$.

So it can be proven that

$$
\begin{aligned}
&|\psi^{(L,\infty)}(\vec{\sigma}_0)/\psi^{(L,\infty)}(\vec{\sigma}')| \\
&\geq \prod_{\tilde{k}=1}^{\tilde{k}_0} \Big[\frac{\cosh((L+1)\mu_0\lambda(\tilde{k}))}{\cosh((L-1)\mu_0\lambda(\tilde{k}))}\Big]^L \times \prod_{\tilde{k}=\tilde{k}_0+1}^{\infty} \Big[\exp(2\beta_3(L-1)\mu_0^2\lambda^2(\tilde{k}))\Big]^L \qquad \text{(A.147)} \\
&\geq \exp\big(2\beta_3 L(L-1)\mu_0^2 \sum_{\tilde{k}=\tilde{k}_0+1}^{\infty} \lambda^2(\tilde{k})\big). \qquad \text{(A.148)}
\end{aligned}
$$

The above lower bound contains infinitely many terms, though each of which is upper bounded by an expression associated with x_0, and can reach arbitrarily high values by decreasing the decaying rate of $\lambda(\tilde{k})$. Therefore, the long-range nature of these RBMs allows the shape of $\psi^{(L,\infty)}(\vec{\sigma})$ in the spin-configuration space to approach the Kronecker delta function as $\lambda(\tilde{k})$ decays more and more slowly. Our numerical results in Fig. A.1 with $\lambda(\tilde{k}) = \delta_P^{-(\tilde{k}-1)}$ imply that the distribution of the square of normalized wave-function amplitudes in the spin-configuration space can approximate the Kronecker delta function with increasing accuracy as $1/\delta_P$ grows up and the ratio $|\psi^{(L,\infty)}(\vec{\sigma}_0)/\psi^{(L,\infty)}(\vec{\sigma}_1)|^2$ as a measure of the approximation accuracy can reach arbitrarily high values as $1/\delta_P$ approaches 1, thus supporting our argument.

A.5 Error curves

In this section, we provide the approximation errors as a function of the number of hidden nodes for truncated LRFD RBMs (Fig. A.2(a)) and the optimal RBMs as well as the truncation errors for the calculations of spin correlations in the x and z directions as a function of the number of levels kept in the truncated LRFD RBMs (Fig. A.2(b)). In

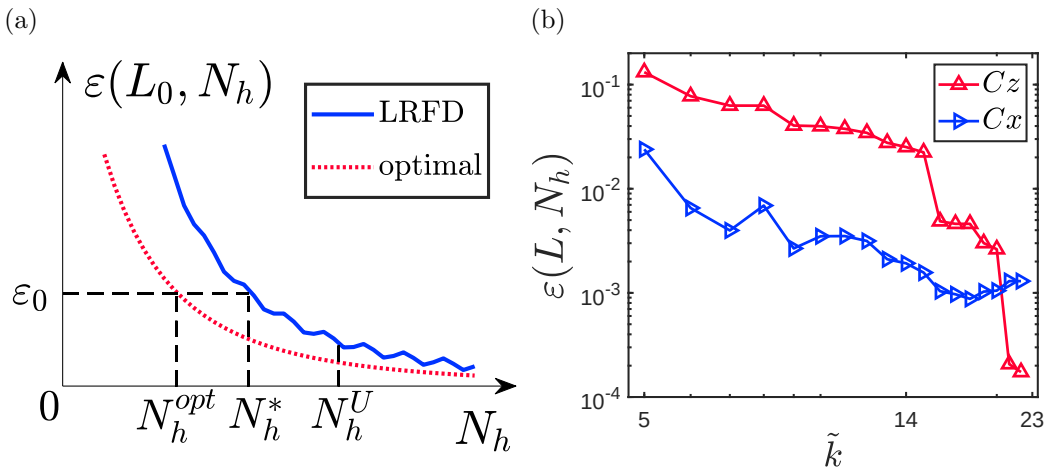


Figure A.2: (a) Approximation errors as a function of the number of hidden nodes for truncated LRFD RBMs and the optimal RBMs. (b) Approximation errors for the calculations of spin correlations in the x and z directions as a function of the number of levels kept in the truncated LRFD RBMs for the XXZ model compared with results from exact-diagonalization methods. The figure is plotted on a log-log scale.

Fig. A.2(a), $\varepsilon(L_0, N_h)$ for LRFD RBMs denotes the second-type truncation errors with $\hat{B}$ generalized into the Hamiltonian of the quantum system in the ground-state learning which usually has a form of a linear combination of polynomially many original $\hat{B}$-type operators. Then the approximation accuracy of using the optimal RBM with the number of hidden nodes not exceeding N_h is definitely better (at least no worse) than that of using an RBM which is a finite truncation of a LRFD RBM keeping N_h hidden nodes based on definitions. So $N_h^U(L_0, \varepsilon_0)$ in our complexity analysis actually also provides an upper bound on $N_h^{\text{opt}}(L_0, \varepsilon_0)$ defined as the minimum number of hidden nodes to achieve a specific approximation error ε_0 for any kinds of RBMs (not limited to LRFD RBMs), which is of great importance for pre-training computational-resource estimations in ML tasks. Fig. A.2(b) shows that the decaying curves of the second-type truncation errors for the ground-state learning of the XXZ model with the same parameter setting as in Fig. 2.4(b) are consistent with our analysis and can be upper bounded by power-law decaying curves

and reach high accuracy as the number of preserved levels increases.

Appendix B: Appendix for Chapter "Learning quasiparticle excitations in long-range interacting quantum systems using neural networks"

B.1 Analysis of energy-shift method

In this appendix, we provide analysis of our energy-shift method, including its implementation, comparison with the orthogonalization method [34], complexity analysis and generalization.

First, we clarify how to implement the projection operators in the energy-shift method. In the stochastic reconfiguration method [6] used in this work, we define the variational log-derivatives with respect to the j-th RBM parameter $\mathcal{W}_j \in \{a_m, b_n, W_{mn} : m = 1, \ldots, L, n = 1, \ldots, N_h\}$ as $O_j(\vec{\sigma}) = \frac{1}{\psi(\vec{\sigma})}\partial_{\mathcal{W}_j}\psi(\vec{\sigma})$ and the so-called local energy as

$$E_{\mathrm{loc}}(\vec{\sigma}) = \langle\vec{\sigma}|\hat{H}|\Psi\rangle/\psi(\vec{\sigma}). \tag{B.1}$$

Then the covariance matrix (also called quantum Fisher matrix [95]) analogous to the Hessian of the energy is $S_{jj'} = \langle O_j^\dagger O_{j'}\rangle - \langle O_j^\dagger\rangle\langle O_{j'}\rangle$ and the force related to the energy gradient is $F_j = \langle O_j^\dagger E_{\mathrm{loc}}\rangle - \langle O_j^\dagger\rangle\langle E_{\mathrm{loc}}\rangle$, where $\langle\ldots\rangle$ means the average over the current

quantum state in the variation process. Then the update of the current parameter vector $\mathcal{W}$ is $-\gamma S^{-1}F$, where γ is a scaling parameter also called the learning rate. Taking the calculation of the first excited state by the energy-shift method as an example, the modification of Hamiltonian in the form of Eq. (4.3) with $j = 1$ only influences the local energy $E_{\text{loc}}(\vec{\sigma})$ by adding a term

$$E_{\text{shift}} \frac{\langle \Psi_0 | \Phi_1 \rangle}{\langle \Psi_0 | \Psi_0 \rangle} \frac{\psi_0(\vec{\sigma})}{\phi_1(\vec{\sigma})}, \tag{B.2}$$

where Φ_1 is the current variational wave function for the first excited state, and $\psi_0(\vec{\sigma})$ and $\phi_1(\vec{\sigma})$ are the amplitudes of spin configuration $\vec{\sigma}$ for $|\Psi_0\rangle$ and $|\Phi_1\rangle$, respectively. The factor $\lambda = \frac{\langle \Psi_0 | \Phi_1 \rangle}{\langle \Psi_0 | \Psi_0 \rangle}$ can be stochastically calculated by

$$\lambda = \sum_{\vec{\sigma}} \frac{|\psi_0(\vec{\sigma})|^2}{\langle \Psi_0 | \Psi_0 \rangle} \left(\frac{\phi_1(\vec{\sigma})}{\psi_0(\vec{\sigma})} \right) \approx \left\langle \frac{\phi_1(\vec{\sigma})}{\psi_0(\vec{\sigma})} \right\rangle_{|\Psi_0\rangle}, \tag{B.3}$$

where $\langle \ldots \rangle_{|\Psi_0\rangle}$ means that the average is carried over samples $\vec{\sigma}$ generated from the probability distribution $|\psi_0(\vec{\sigma})|^2$ through Monte-Carlo sampling.

Second, our energy-shift method can be compared with the orthogonalization method for resolving excited states without using translational symmetry information [34]. That orthogonalization method calculates the first-excited state by eliminating the overlap of $|\Phi_1\rangle$ with the ground state $|\Psi_0\rangle$ at every iteration step during the variation process. Both methods need stochastic calculations over two probability distributions $|\psi_0(\vec{\sigma})|^2$ and $|\phi_1(\vec{\sigma})|^2$. But our energy-shift method only needs to modify the calculation of F at every iteration step while the orthogonalization method needs to modify both S and F as the

wave function itself is changed into a difference between two RBMs: $|\Phi_1\rangle - \lambda|\Psi_0\rangle$. Though λ becomes very small at the end of variation process, our energy-shift method has an advantage that the target excited state is represented by a single RBM, rather than by multiple RBMs, throughout the entire optimization path. Such a single-RBM representation is more natural and has a smaller parameter space, so the information geometry, such as the quantum Fisher matrix interpreted as the Riemannian metric tensor, along the optimization path can be investigated more easily, which will help understand the corresponding quantum phase [94, 95, 96]. It is also worthwhile to study how the energy-shift idea implicitly changes the pathological curvature in the network parameter space [94] and influences the convergence rate and numerical stability of the optimization algorithm.

Third, our training data shows that the quantum Fisher matrices S for long-range interacting systems often have larger condition numbers, the ratio of the maximal to the minimal singular values of S, compared to short-range interacting systems in the same phase. Let N_{par} denote the number of variational parameters which is equal to $L+N_h+LN_h$ for an RBM, and N_{MC} denote the number of Monte Carlo sweeps. As the condition number of S is associated with the structure of the eigenspace of S [124], this feature leads to that the number of vectors in Krylov subspace required to efficiently approximate the solution $S^{-1}F$ increases from $\mathcal{O}(1)$ to $\mathcal{O}(N_{\text{par}})$ when we use the iterative Krylov subspace solver and conjugate-gradient descent [6] for parameter update. So the computational complexity for solving $S^{-1}F$ at every iteration, which is the most time-consuming part of the stochastic reconfiguration method, increases from $\mathcal{O}(N_{\text{par}}N_{\text{MC}})$ to $\mathcal{O}(N_{\text{par}}^2N_{\text{MC}})$. This result implies that the stronger correlations and possibly larger entanglements in the energy eigenstates of long-range interacting systems may lead to less steep learning landscape [95] and thus

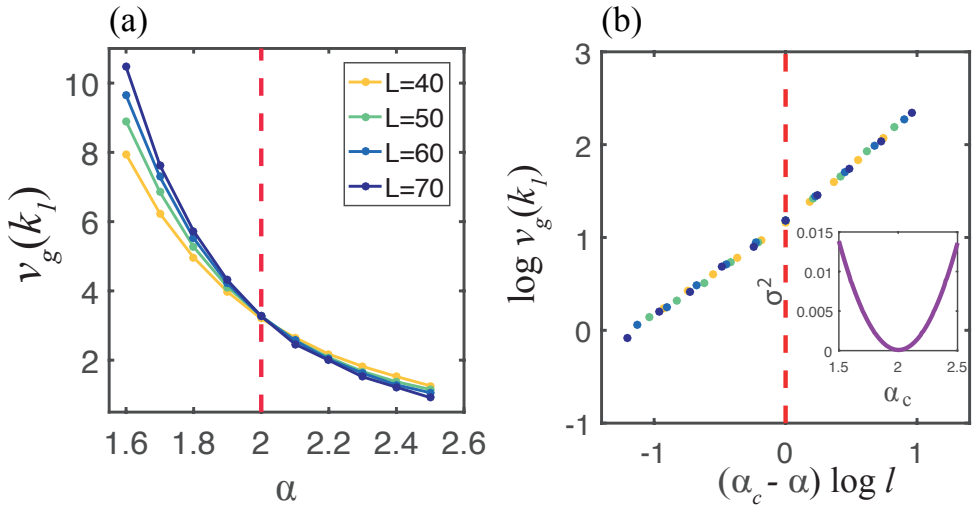


Figure B.1: Long-range TFIM with $B = 6$ learned with an RBM with hidden-unit density equal to 1. (a) Group velocity $v_g(k_1)$ as a function of α for a range of system sizes L. (b) Data collapse for (a), while the inset shows the variance in data collapse processing for different choices of α_c.

slower learning rates compared to short-range interacting systems.

At last, the incorporation of our energy-shift method and the fixed-quantum-number method can be used to calculate not only the lowest, but also the second (third,...) lowest excited states in each momentum sector. For example, suppose $|\Psi_{k,1}\rangle$ and $|\Psi_{k,2}\rangle$ are the first and second excited states in the sector of momentum k. To get $|\Psi_{k,2}\rangle$ with obtained $|\Psi_{k,1}\rangle$, we just need to construct a wave function ansatz still satisfying the translational symmetry constraints associated with k and add $E^{(\text{shift})}\frac{|\Psi_{k,1}\rangle\langle\Psi_{k,1}|}{\langle\Psi_{k,1}|\Psi_{k,1}\rangle}$ to the Hamiltonian to shift the energy of $|\Psi_{k,1}\rangle$ upward making $|\Psi_{k,2}\rangle$ the new ground state in this sector. The main limitation of this idea will be the expressibility of an RBM itself for slightly higher excited states.

B.2 Field-theory formula analysis and data collapse

Based on the field-theoretic dispersion-relation formula in the long-wavelength limit (Eq. (4.5) in the main text) [97], we find that, as $k \to 0$:

for the gapped TFIM,

$$\omega(k) \approx \Delta[1 + \frac{1}{2\Delta^2}(ak^{\alpha-1} + bk^2)], \tag{B.4}$$

$$v_g(k) = \frac{d\omega(k)}{dk} \approx \frac{a}{2\Delta}(\alpha - 1)k^{\alpha-2} + \frac{b}{\Delta}k; \tag{B.5}$$

for the gapless XXZ model,

$$\omega(k) \approx (ak^{\alpha-1} + bk^2)^{\frac{1}{2}}, \tag{B.6}$$

$$v_g(k) \approx \begin{cases} a^{\frac{1}{2}}(\frac{\alpha-1}{2})k^{\frac{\alpha-3}{2}} + \frac{b}{2a^{\frac{1}{2}}}(\frac{5-\alpha}{2})k^{\frac{3-\alpha}{2}}, \text{ for } 1 < \alpha < 3, \\ b^{\frac{1}{2}} + \frac{a}{2b^{\frac{1}{2}}}(\alpha - 2)k^{\alpha-3}, \text{ for } \alpha > 3. \end{cases} \tag{B.7}$$

So the monotonic behavior of $v_g(k_1)$ with growing L (i.e., decreasing k_1) across α_c in this analysis are consistent with our numerical results.

We now provide more details behind the data-collapse method used in the main text to compute the critical exponent α_c. This method is useful since identifying α_c directly by the sign-flip of the dispersion curvature near $k = 0$ suffers from large finite-size errors. To quantitatively measure the collapsing quality [104] with different choices of α_c, we compute the variance σ^2, which sums over the deviation of data points from the scaling function.

We calculate the variance for a specific α_c as follows:

$$\sigma^2 = \frac{1}{n-1}\sum_{j=1}^{n} \mathrm{Var}(f(x_j)), \tag{B.8}$$

where for the TFIM, all the values of $(\alpha_c - \alpha)\log l$ for a range of L and α form a set $X = \{x_j, 1 \leq j \leq n\}$, n is the number of unique elements in X, and each x_j and its corresponding $(\log v_g(k_1))_j$ form a data point. Since the points at which the scaling function $f(x)$ is evaluated are different for each system size, we estimate the value of $f(x)$ for each L by linearly extrapolating the data only from that L. Thus, we can get the variance $\mathrm{Var}(f(x_j))$ at each x_j and then do an average.

Figs. B.1(a) and (b) show the behavior of group velocity $v_g(k_1)$ and the data collapse, respectively, for the TFIM with $B = 6$, as compared to $B = 4$ in the main text. The qualitatively same results for both $B = 4$ and $B = 6$ support the correctness of our analysis for the TFIM in the paramagnetic phase. It is remarkable that the borderline values of the power-law exponent $\alpha = 1$ or α_c often yield a multiplicative logarithm in the dispersions and may bring about new types of phases and phase transitions [68]. Therefore, these two cases deserve future study.

Appendix C: Appendix for Chapter "Neural networks as phase classifiers"

C.1 Hamiltonian in Majorana representation

Here we show our Hamiltonian $\hat{H}$ expressed as a sum of products of Majorana operators, γ. The generalized "sweet spot" conditions are apparent in this representation. In our model, we choose to make the terms of $\gamma_{\vec{r}_j,\uparrow,1}\gamma_{\vec{r}_{j'},\uparrow,1}$, $\gamma_{\vec{r}_j,\uparrow,1}\gamma_{\vec{r}_{j'},\uparrow,2}$, $\gamma_{\vec{r}_j,\downarrow,2}\gamma_{\vec{r}_{j'},\downarrow,2}$ and $\gamma_{\vec{r}_j,\downarrow,2}\gamma_{\vec{r}_{j'},\downarrow,1}$ vanish, which will lead to Eqs.(4.4)–(4.6).

$$\begin{aligned}\hat{H} = \frac{i}{2}\sum_{\langle j,m\rangle}\Big[(-t-\Delta)\gamma_{\vec{r}_j,\uparrow,1}\gamma_{\vec{r}_m,\downarrow,2} + (t-\Delta)\gamma_{\vec{r}_j,\uparrow,2}\gamma_{\vec{r}_m,\downarrow,1}\Big] \\ +\frac{i}{2}\sum_{\langle\langle j,j'\rangle\rangle,y_j<y_{j'}}\sum_{\sigma=\uparrow,\downarrow}\Big[(-t_\sigma\sin\phi_\sigma + \mathrm{Im}(\Delta_\sigma))\gamma_{\vec{r}_j,\sigma,1}\gamma_{\vec{r}_{j'},\sigma,1} \\ +(-t_\sigma\sin\phi_\sigma - \mathrm{Im}(\Delta_\sigma))\gamma_{\vec{r}_j,\sigma,2}\gamma_{\vec{r}_{j'},\sigma,2} \\ +(-t_\sigma\cos\phi_\sigma - \mathrm{Re}(\Delta_\sigma))\gamma_{\vec{r}_j,\sigma,1}\gamma_{\vec{r}_{j'},\sigma,2} \\ +(t_\sigma\cos\phi_\sigma - \mathrm{Re}(\Delta_\sigma))\gamma_{\vec{r}_j,\sigma,2}\gamma_{\vec{r}_{j'},\sigma,1}\Big]. \end{aligned} \tag{C.1}$$

C.2 Derivation of the phase boundary between gapped SC and gapless phases

This section derives the expression for the phase boundary between gapped SC and gapless phases in the $p_x - p_y$ plane when other parameters in our model are fixed. The results are in close agreement with the phase diagram (Fig. 4.4(a)) that was obtained by numerical simulations of band structures for every point in $p_x - p_y$ plane. This supports the correctness of our mathematical analysis, exemplifies the bulk-edge correspondence in topological physics and provides the basis for our identification of "strong vs. weak" of TRS breaking.

According to the results of section IV, the band touching condition reduces to the vanishing of the discriminant given in Eq. (4.10). We now show the steps of deriving the inequality (4.11).

Let $z_j = e^{i\vec{k}\cdot\vec{r}_j}$, $a_j = \cos(\vec{p}\cdot\vec{r}_j/2)$, for $j = 1, 2, 3$. Note that $z_2 = z_1 z_3$ since $\vec{r}_2 = \vec{r}_1 + \vec{r}_3$. For any (z_1, z_3) such that $|z_1| = |z_3| = 1$, the corresponding vector $\vec{k} = (k_x, k_y)$ can be determined. Thus, we have converted the problem of finding such a $\vec{k}$ to a problem of finding z_j. Using "$\Leftrightarrow$" to denote equivalence, it is easy to show that:

$$\sum_{j=1}^{3} z_j a_j = 0 \Leftrightarrow z_1 = \frac{-a_3}{a_1 z_3^{-1} + a_2}.$$

$$\text{So } |z_j| = 1 (j = 1, 2, 3) \Leftrightarrow \quad \text{there exists a } z_3 \text{ with } |z_3| = 1$$
$$\text{such that } |a_3| = |a_1 z_3^{-1} + a_2|.$$

Considering that $||a_1| - |a_2|| \leq |a_1 z_3^{-1} + a_2| \leq |a_1| + |a_2|$ for $|z_3| = 1$, the above requirement

$$\Leftrightarrow ||a_1| - |a_2|| \leq |a_3| \leq |a_1| + |a_2|$$
$$\Leftrightarrow |a_1| - |a_2| \leq |a_3|, |a_2| - |a_1| \leq |a_3|, |a_3| \leq |a_1| + |a_2|.$$

Therefore, the band touching condition is finally reduced to:

$$|a_j| + |a_k| \geq |a_l|,$$

for any (j, k, l) being a permutation of $(1, 2, 3)$. This is the inequality (4.11). Any special case with singularities can be verified to satisfy these conditions.

Appendix D: Other research projects

D.1 Optomechanics and novel quantum phase of the Bose-Einstein Condensate with the cavity mediated spin-orbit coupling

During my Ph.D. study, I also conducted research on several other research topics, such as "optomechanics and novel quantum phase of the Bose-Einstein Condensate with the cavity mediated spin-orbit coupling" and "multiple photon-echo rephasing of coherent matter waves" [125, 126]. In this section, I will briefly introduce the former research project in which I cooperated with other researchers at the Peking University in China. It is not included in the subject of quantum machine learning, but belongs to general atomic, molecular and optical physics.

Abstract:

We investigated the optomechanical dynamics and explored the quantum phase of a Bose-Einstein condensate in a ring cavity. The interaction between the atoms and the cavity field in the running wave mode induces a semiquantized spin-orbit coupling (SOC) for the atoms. We found that the evolution of the magnetic excitations of the matter field resembles that of an optomechanical oscillator moving in a viscous optical medium, with

very good integrability and traceability, regardless of the atomic interaction. Moreover, the light-atom coupling induces a sign-changeable long-range interatomic interaction, which reshapes the typical energy spectrum of the system in a drastic manner. As a result, a new quantum phase featured a high quantum degeneracy was found in the transitional area for SOC. Our scheme is immediately realizable and the results are measurable in experiments.

Introduction:

Cavity quantum electrodynamics (CQED) describes the coherent interaction between matter fields and electromagnetic (EM) fields within optical resonators. In the CQED systems, the dynamics of the matter field and that of the EM field greatly affect each other. Cavity optomechanics is one of the good examples studying the coupling between a quantum mechanical oscillator and the EM field inside a cavity. It has been a continual refreshing subject in the cold atom physics. Through the combination of CQED and the cold atom physics, a wide range of exotic quantum phenomena have been investigated. In the past few years, there has been a breakthrough of realizing the spin-orbit coupling (SOC) in cold atom gases which could be naturally incorporated into the CQED systems. Perspectives of new quantum kinetics and fascinating phenomena are being opened up in the relevant areas like realizing quantum system with synthetic dimensions, ideal Weyl semimetal, and superradiant phase transition.

In the experiments of the ultracold atomic gases with the SOC in open space, the coupling between the atomic spin (or pseudo spin) and the center of mass (c.m.) motion is induced by utilizing the classical Raman laser beams that are far-detuned with the atomic resonance frequencies. Then the coupling between lights and atoms is weak; thus the

fluctuation of the light fields is negligible. However, when the atoms are placed inside the cavity, the atoms can exert a strong influence on the cavity field; thus strong nonlinear effects would arise. Meanwhile, because all the atoms couple to the same cavity field, long-range atomic correlation could be induced. With the help of optical resonators, different kinds of cavity-mediated long-range atomic interactions have been realized in investigations on dynamical phase transitions, spin entanglement, superradiant quantum phase transition, etc. Such nonlinearity and long-range correlation generate many intriguing effects that are unique in the CQED systems.

This paper is devoted to study the optomechanics and rich quantum phase of a Bose-Einstein condensate (BEC) contained in a ring cavity in which a light in the running wave mode is maintained. It is a typical CQED system that can give rise to a semiquantized SOC for the atoms. We are going to show that this system is an optomechanical system that the optomechanical kinetics of the matter wave excitations is highly integrable and traceable. This is an valuable merit in contrast to the previous works. Moreover, it has been demonstrated that CQED systems with SOC can have very rich quantum phases and complex dispersion relations; and along with these phenomena, there is the typical double minima to single-minimum transition in the energy dispersion relation. However, due to the presence of the long-range correlation among the atoms induced by the SOC in the cavity, a new quantum phase with high quantum degeneracy will appear at the transition point. The latter part of this work is to uncover this new phase. This paper is organized as the following: First, we present the principal model and theory of the CQED system with the SOC. Second, we study the properties of the optomechanical motion of the magnetic excitations of the system. Third, we demonstrate a detailed analysis of the spectrum of

the elementary excitations, and present the phase diagram of the magnetic order of the matter field, in which the new quantum phase is included. Finally, we give a discussion and summary of our results.

Discussion and conclusion:

We would like to point out that the damping optomechanical motion of the elementary excitations of the CQED system is intrinsic. The physical observables, such as the number of the photons inside the cavity and the atomic occupations on the spin-up and -down states, gradually reach the steady stage after a few periods of oscillations. This is very profitable for preparing a BEC in the thermal equilibrium in a real experiment by using the radio frequency technique, lattice shaking, etc. In a trap, there are discrete vibrational excited states, thus it is very necessary that the atoms be firstly cooled down to condense in the ground state of the trap to form a BEC by using cooling techniques in order to avoid the excited states being involved. Then one can turn on the lights in the CQED system so that the atoms interact with the light fields. This induces the cavity-mediated SOC and the long-range atomic interaction. Finally, by using the above artificial thermalization methods and cooling the atoms down again at the same time, we could arrive at the true ground state for the whole CQED system, and observe the quantum phase transitions in the corresponding parameter region.

The setup under our consideration is a very typical configuration of a large group of CQED systems with SOC. In these systems, quantum phase transitions may take place at the double-minima to single-minimum transition of the dispersion relation under the corresponding parameter conditions. However, our work shows that, right at the transition

point (or region), there is a new quantum phase (the multi-fold degeneracy (MD) phase) with a higher quantum degeneracy and multi-fold (usually more than two folds) momentum peaks. This fills the lack of relevant investigations of quantum phase transitions in the transitional area. It would add fine structures to the complete quantum phase diagrams for the CQED systems with SOC, and it would help us to dig into the salient features of the transiting behaviors of the quantum orders in these CQED systems.

In conclusion, we have studied the optomechanics and the novel quantum phases of a BEC in the CQED system, which utilizes the traveling light inside a ring cavity to induce the semiquantized SOC and a highly tunable long-range atomic interaction in the matter field. We found that the motion of the magnetic excitations in the matter field can be mapped perfectly onto that of an optomechanical oscillator moving in the viscous optical medium. This optomechanical motion possesses high and exact traceability which is exempt from the influence of atomic interactions. It may find its usage in the high-precision measurements such as atomic clocks, and extensions to other hybrid systems. In this system, the interrelatedness between the cavity field and the matter field is strong and nonlinear that the spectrum of the elementary excitations can be shaped into forms with diverse quantum degeneracies, especially in the transitional parameter region. This results in a new quantum phase–the MD phase which deserves our attention but has never been brought into light before. The notions and results of this work are readily realizable in experiments. They could enable new investigations in quantum optics, many-body physics, and quantum magnetism.

Bibliography

[1] Giuseppe Carleo, Ignacio Cirac, Kyle Cranmer, Laurent Daudet, Maria Schuld, Naftali Tishby, Leslie Vogt-Maranto, and Lenka Zdeborová. Machine learning and the physical sciences. *Rev. Mod. Phys.*, 91:045002, Dec 2019.

[2] Jacob Biamonte, Peter Wittek, Nicola Pancotti, Patrick Rebentrost, Nathan Wiebe, and Seth Lloyd. Quantum machine learning. *Nature*, 549(7671):195–202, 2017.

[3] Juan Carrasquilla and Roger G Melko. Machine learning phases of matter. *Nat. Phys.*, 13(5):431–434, 2017.

[4] Hsin-Yuan Huang, Richard Kueng, Giacomo Torlai, Victor V. Albert, and John Preskill. Provably efficient machine learning for quantum many-body problems. *Science*, 377(6613):eabk3333, 2022.

[5] Sankar Das Sarma, Dong-Ling Deng, and Lu-Ming Duan. Machine learning meets quantum physics. *Phys. Today*, 72(3):48, Mar 2019.

[6] Giuseppe Carleo and Matthias Troyer. Solving the quantum many-body problem with artificial neural networks. *Science*, 355(6325):602–606, 2017.

[7] Roger G Melko, Giuseppe Carleo, Juan Carrasquilla, and J Ignacio Cirac. Restricted boltzmann machines in quantum physics. *Nat. Phys.*, 15(9):887–892, June 2019.

[8] Ivan Glasser, Nicola Pancotti, Moritz August, Ivan D. Rodriguez, and J. Ignacio Cirac. Neural-network quantum states, string-bond states, and chiral topological states. *Phys. Rev. X*, 8:011006, Jan 2018.

[9] Xun Gao and Lu-Ming Duan. Efficient representation of quantum many-body states with deep neural networks. *Nat. Commun.*, 8:662, September 2017.

[10] Giacomo Torlai and Roger G. Melko. Learning thermodynamics with boltzmann machines. *Phys. Rev. B*, 94:165134, Oct 2016.

[11] Seth Lloyd, Masoud Mohseni, and Patrick Rebentrost. Quantum algorithms for supervised and unsupervised machine learning. *arXiv preprint arXiv:1307.0411*, 2013.

[12] Nathan Wiebe, Ashish Kapoor, and Krysta M. Svore. Quantum algorithms for nearest-neighbor methods for supervised and unsupervised learning. *Quantum Info. Comput.*, 15(3–4):316, March 2015.

[13] Nathan Wiebe, Ashish Kapoor, and Krysta M Svore. Quantum deep learning. *arXiv preprint arXiv:1412.3489*, 2014.

[14] Patrick Rebentrost, Masoud Mohseni, and Seth Lloyd. Quantum support vector machine for big data classification. *Phys. Rev. Lett.*, 113:130503, Sep 2014.

[15] Mohammad H. Amin, Evgeny Andriyash, Jason Rolfe, Bohdan Kulchytskyy, and Roger Melko. Quantum boltzmann machine. *Phys. Rev. X*, 8:021050, May 2018.

[16] Pankaj Mehta and David J Schwab. An exact mapping between the variational renormalization group and deep learning. *arXiv preprint arXiv:1410.3831*, 2014.

[17] Wenjian Hu, Rajiv R. P. Singh, and Richard T. Scalettar. Discovering phases, phase transitions, and crossovers through unsupervised machine learning: A critical examination. *Phys. Rev. E*, 95:062122, Jun 2017.

[18] Kelvin Ch'ng, Juan Carrasquilla, Roger G. Melko, and Ehsan Khatami. Machine learning phases of strongly correlated fermions. *Phys. Rev. X*, 7:031038, Aug 2017.

[19] Lei Wang. Discovering phase transitions with unsupervised learning. *Phys. Rev. B*, 94:195105, Nov 2016.

[20] Ruizhi Pan and Charles W. Clark. Implementing majorana fermions in a cold-atom honeycomb lattice with textured pairings. *Phys. Rev. A*, 98:033604, Sep 2018.

[21] Yi-Ting Hsu, Xiao Li, Dong-Ling Deng, and S. Das Sarma. Machine learning many-body localization: Search for the elusive nonergodic metal. *Phys. Rev. Lett.*, 121:245701, Dec 2018.

[22] Matthias Rupp, Alexandre Tkatchenko, Klaus-Robert Müller, and O. Anatole von Lilienfeld. Fast and accurate modeling of molecular atomization energies with machine learning. *Phys. Rev. Lett.*, 108:058301, Jan 2012.

[23] Albert P Bartók, Sandip De, Carl Poelking, Noam Bernstein, James R Kermode, Gábor Csányi, and Michele Ceriotti. Machine learning unifies the modeling of materials and molecules. *Sci. Adv.*, 3(12):e1701816, 2017.

[24] Giacomo Torlai, Guglielmo Mazzola, Juan Carrasquilla, Matthias Troyer, Roger Melko, and Giuseppe Carleo. Neural-network quantum state tomography. *Nat. Phys.*, 14(5):447, 2018.

[25] Mária Kieferová and Nathan Wiebe. Tomography and generative training with quantum boltzmann machines. *Phys. Rev. A*, 96:062327, Dec 2017.

[26] Junwei Liu, Yang Qi, Zi Yang Meng, and Liang Fu. Self-learning monte carlo method. *Phys. Rev. B*, 95:041101, Jan 2017.

[27] Li Huang and Lei Wang. Accelerated monte carlo simulations with restricted boltzmann machines. *Phys. Rev. B*, 95:035105, Jan 2017.

[28] Zi Cai and Jinguo Liu. Approximating quantum many-body wave functions using artificial neural networks. *Phys. Rev. B*, 97(3):035116, 2018.

[29] Yusuke Nomura, Andrew S. Darmawan, Youhei Yamaji, and Masatoshi Imada. Restricted boltzmann machine learning for solving strongly correlated quantum systems. *Phys. Rev. B*, 96:205152, Nov 2017.

[30] Dong-Ling Deng, Xiaopeng Li, and S. Das Sarma. Machine learning topological states. *Phys. Rev. B*, 96:195145, Nov 2017.

[31] Dong-Ling Deng, Xiaopeng Li, and S. Das Sarma. Quantum entanglement in neural network states. *Phys. Rev. X*, 7:021021, May 2017.

[32] J. Ignacio Cirac, David Pérez-García, Norbert Schuch, and Frank Verstraete. Matrix product states and projected entangled pair states: Concepts, symmetries, theorems. *Rev. Mod. Phys.*, 93:045003, Dec 2021.

[33] Nicolas Le Roux and Yoshua Bengio. Representational power of restricted boltzmann machines and deep belief networks. *Neural Comput.*, 20(6):1631–1649, 2008.

[34] Kenny Choo, Giuseppe Carleo, Nicolas Regnault, and Titus Neupert. Symmetries and many-body excitations with neural-network quantum states. *Phys. Rev. Lett.*, 121:167204, Oct 2018.

[35] Yunqin Zheng, Huan He, Nicolas Regnault, and B. Andrei Bernevig. Restricted boltzmann machines and matrix product states of one-dimensional translationally invariant stabilizer codes. *Phys. Rev. B*, 99:155129, Apr 2019.

[36] Huan He, Yunqin Zheng, B Andrei Bernevig, and German Sierra. Multi-layer restricted boltzmann machine representation of 1d quantum many-body wave functions. *arXiv preprint arXiv:1910.13454*, 2019.

[37] Raphael Kaubruegger, Lorenzo Pastori, and Jan Carl Budich. Chiral topological phases from artificial neural networks. *Phys. Rev. B*, 97:195136, May 2018.

[38] Yichen Huang and Joel E. Moore. Neural network representation of tensor network and chiral states. *Phys. Rev. Lett.*, 127:170601, Oct 2021.

[39] M. Saffman, T. G. Walker, and K. Mølmer. Quantum information with rydberg atoms. *Rev. Mod. Phys.*, 82:2313–2363, Aug 2010.

[40] Bo Yan, Steven A Moses, Bryce Gadway, Jacob P Covey, Kaden RA Hazzard, Ana Maria Rey, Deborah S Jin, and Jun Ye. Observation of dipolar spin-exchange interactions with lattice-confined polar molecules. *Nature*, 501(7468):521–525, 2013.

[41] K. Aikawa, A. Frisch, M. Mark, S. Baier, A. Rietzler, R. Grimm, and F. Ferlaino. Bose-einstein condensation of erbium. *Phys. Rev. Lett.*, 108:210401, May 2012.

[42] Mingwu Lu, Nathaniel Q. Burdick, and Benjamin L. Lev. Quantum degenerate dipolar fermi gas. *Phys. Rev. Lett.*, 108:215301, May 2012.

[43] Florian Dolde, Ingmar Jakobi, Boris Naydenov, Nan Zhao, Sebastien Pezzagna, Christina Trautmann, Jan Meijer, Philipp Neumann, Fedor Jelezko, and Jörg Wrachtrup. Room-temperature entanglement between single defect spins in diamond. *Nat. Phys.*, 9(3):139, 2013.

[44] Peter Schauß, Marc Cheneau, Manuel Endres, Takeshi Fukuhara, Sebastian Hild, Ahmed Omran, Thomas Pohl, Christian Gross, Stefan Kuhr, and Immanuel Bloch. Observation of spatially ordered structures in a two-dimensional rydberg gas. *Nature*, 491(7422):87–91, 2012.

[45] Johannes Zeiher, Rick Van Bijnen, Peter Schauß, Sebastian Hild, Jae-yoon Choi, Thomas Pohl, Immanuel Bloch, and Christian Gross. Many-body interferometry of a rydberg-dressed spin lattice. *Nat. Phys.*, 12(12):1095, Aug 2016.

[46] Ofer Firstenberg, Thibault Peyronel, Qi-Yu Liang, Alexey V Gorshkov, Mikhail D Lukin, and Vladan Vuletić. Attractive photons in a quantum nonlinear medium. *Nature*, 502(7469):71–75, 2013.

[47] R. Islam, C. Senko, W. C. Campbell, S. Korenblit, J. Smith, A. Lee, E. E. Edwards, C.-C. J. Wang, J. K. Freericks, and C. Monroe. Emergence and frustration of magnetism with variable-range interactions in a quantum simulator. *Science*, 340(6132):583, May 2013.

[48] Philip Richerme, Zhe-Xuan Gong, Aaron Lee, Crystal Senko, Jacob Smith, Michael Foss-Feig, Spyridon Michalakis, Alexey V Gorshkov, and Christopher Monroe. Non-local propagation of correlations in quantum systems with long-range interactions. *Nature*, 511(7508):198–201, 2014.

[49] Petar Jurcevic, Ben P Lanyon, Philipp Hauke, Cornelius Hempel, Peter Zoller, Rainer Blatt, and Christian F Roos. Quasiparticle engineering and entanglement propagation in a quantum many-body system. *Nature*, 511(7508):202–205, 2014.

[50] Joseph W Britton, Brian C Sawyer, Adam C Keith, C-C Joseph Wang, James K Freericks, Hermann Uys, Michael J Biercuk, and John J Bollinger. Engineered two-dimensional ising interactions in a trapped-ion quantum simulator with hundreds of spins. *Nature*, 484(7395):489–492, 2012.

[51] J. G. Bohnet, B. C. Sawyer, J. W. Britton, M. L. Wall, A. M. Rey, M. Foss-Feig, and J. J. Bollinger. Quantum spin dynamics and entanglement generation with hundreds of trapped ions. *Science*, 352(6291):1297, Jun 2016.

[52] C. Monroe, W. C. Campbell, L.-M. Duan, Z.-X. Gong, A. V. Gorshkov, P. W. Hess, R. Islam, K. Kim, N. M. Linke, G. Pagano, P. Richerme, C. Senko, and N. Y. Yao. Programmable quantum simulations of spin systems with trapped ions. *Rev. Mod. Phys.*, 93:025001, Apr 2021.

[53] Anna Golubeva and Roger G. Melko. Pruning a restricted boltzmann machine for quantum state reconstruction. *Phys. Rev. B*, 105:125124, Mar 2022.

[54] Dan Sehayek, Anna Golubeva, Michael S. Albergo, Bohdan Kulchytskyy, Giacomo Torlai, and Roger G. Melko. Learnability scaling of quantum states: Restricted boltzmann machines. *Phys. Rev. B*, 100:195125, Nov 2019.

[55] Stephen R Clark. Unifying neural-network quantum states and correlator product states via tensor networks. *J. Phys. A Math. Theor.*, 51(13):135301, feb 2018.

[56] Jing Chen, Song Cheng, Haidong Xie, Lei Wang, and Tao Xiang. Equivalence of restricted boltzmann machines and tensor network states. *Phys. Rev. B*, 97:085104, Feb 2018.

[57] J. Eisert, M. Cramer, and M. B. Plenio. Colloquium: Area laws for the entanglement entropy. *Rev. Mod. Phys.*, 82:277–306, Feb 2010.

[58] Zhoushen Huang and Alexander V Balatsky. Complexity and geometry of quantum state manifolds. *arXiv preprint arXiv:1711.10471*, 2017.

[59] Mikito Koshino, Noah F. Q. Yuan, Takashi Koretsune, Masayuki Ochi, Kazuhiko Kuroki, and Liang Fu. Maximally localized wannier orbitals and the extended hubbard model for twisted bilayer graphene. *Phys. Rev. X*, 8:031087, Sep 2018.

[60] Roger A Horn and Charles R Johnson. *Matrix analysis*. Cambridge university press, Chap. 5, p. 274, 2012.

[61] Roger A Horn and Charles R Johnson. *Matrix analysis*. Cambridge university press, Chap. 5, p. 265, 2012.

[62] F. Verstraete and J. I. Cirac. Matrix product states represent ground states faithfully. *Phys. Rev. B*, 73:094423, Mar 2006.

[63] D. Perez-Garcia, F. Verstraete, M. M. Wolf, and J. I. Cirac. Matrix product state representations. *Quantum Info. Comput.*, 7(5):401–430, jul 2007.

[64] Sandro Sorella, Michele Casula, and Dario Rocca. Weak binding between two aromatic rings: Feeling the van der waals attraction by quantum monte carlo methods. *J. Chem. Phys.*, 127(1):014105, Jul 2007.

[65] Eric Neuscamman, C. J. Umrigar, and Garnet Kin-Lic Chan. Optimizing large parameter sets in variational quantum monte carlo. *Phys. Rev. B*, 85:045103, Jan 2012.

[66] A. Harju, B. Barbiellini, S. Siljamäki, R. M. Nieminen, and G. Ortiz. Stochastic gradient approximation: An efficient method to optimize many-body wave functions. *Phys. Rev. Lett.*, 79:1173, Aug 1997.

[67] Daniel Jaschke, Kenji Maeda, Joseph D Whalen, Michael L Wall, and Lincoln D Carr. Critical phenomena and kibble–zurek scaling in the long-range quantum ising chain. *New J. Phys.*, 19(3):033032, Mar 2017.

[68] Mohammad F. Maghrebi, Zhe-Xuan Gong, and Alexey V. Gorshkov. Continuous symmetry breaking in 1d long-range interacting quantum systems. *Phys. Rev. Lett.*, 119:023001, Jul 2017.

[69] Thomas Koffel, M. Lewenstein, and Luca Tagliacozzo. Entanglement entropy for the long-range ising chain in a transverse field. *Phys. Rev. Lett.*, 109:267203, Dec 2012.

[70] Pochung Chen, Zhi long Xue, I P McCulloch, Ming-Chiang Chung, Miguel Cazalilla, and S-K Yip. Entanglement entropy scaling of the xxz chain. *J. Stat. Mech. Theory Exp.*, 2013(10):P10007, oct 2013.

[71] Francisco C. Alcaraz and Márcio J. Martins. Conformal invariance and the heisenberg model with arbitrary spin. *Phys. Rev. Lett.*, 63:708–711, Aug 1989.

[72] Yi Zhang and Eun-Ah Kim. Quantum loop topography for machine learning. *Phys. Rev. Lett.*, 118:216401, May 2017.

[73] Giacomo Torlai and Roger G. Melko. Latent space purification via neural density operators. *Phys. Rev. Lett.*, 120:240503, Jun 2018.

[74] Nathan Wiebe and Christopher Granade. Can small quantum systems learn? *arXiv preprint arXiv:1512.03145*, 2015.

[75] Vedran Dunjko, Jacob M. Taylor, and Hans J. Briegel. Quantum-enhanced machine learning. *Phys. Rev. Lett.*, 117:130501, Sep 2016.

[76] P. Hauke and L. Tagliacozzo. Spread of correlations in long-range interacting quantum systems. *Phys. Rev. Lett.*, 111:207202, Nov 2013.

[77] J. Schachenmayer, B. P. Lanyon, C. F. Roos, and A. J. Daley. Entanglement growth in quench dynamics with variable range interactions. *Phys. Rev. X*, 3:031015, Sep 2013.

[78] Elliott H. Lieb and Derek W. Robinson. The finite group velocity of quantum spin systems. *Comm. Math. Phys.*, 28(3):251, 1972.

[79] Laurens Vanderstraeten, Maarten Van Damme, Hans Peter Büchler, and Frank Verstraete. Quasiparticles in quantum spin chains with long-range interactions. *Phys. Rev. Lett.*, 121:090603, Aug 2018.

[80] Anna Posazhennikova, Mauricio Trujillo-Martinez, and Johann Kroha. Inflationary quasiparticle creation and thermalization dynamics in coupled bose-einstein condensates. *Phys. Rev. Lett.*, 116:225304, Jun 2016.

[81] Cheng-Ju Lin and Olexei I. Motrunich. Quasiparticle explanation of the weak-thermalization regime under quench in a nonintegrable quantum spin chain. *Phys. Rev. A*, 95:023621, Feb 2017.

[82] Marton Kormos, Mario Collura, Gabor Takács, and Pasquale Calabrese. Real-time confinement following a quantum quench to a non-integrable model. *Nat. Phys.*, 13(3):246, Nov 2016.

[83] Anton S. Buyskikh, Maurizio Fagotti, Johannes Schachenmayer, Fabian Essler, and Andrew J. Daley. Entanglement growth and correlation spreading with variable-range interactions in spin and fermionic tunneling models. *Phys. Rev. A*, 93:053620, May 2016.

[84] Lorenzo Cevolani, Julien Despres, Giuseppe Carleo, Luca Tagliacozzo, and Laurent Sanchez-Palencia. Universal scaling laws for correlation spreading in quantum systems with short- and long-range interactions. *Phys. Rev. B*, 98:024302, Jul 2018.

[85] Irénée Frérot, Piero Naldesi, and Tommaso Roscilde. Entanglement and fluctuations in the xxz model with power-law interactions. *Phys. Rev. B*, 95:245111, Jun 2017.

[86] D. Peter, S. Müller, S. Wessel, and H. P. Büchler. Anomalous behavior of spin systems with dipolar interactions. *Phys. Rev. Lett.*, 109:025303, Jul 2012.

[87] Zhe-Xuan Gong, Michael Foss-Feig, Spyridon Michalakis, and Alexey V. Gorshkov. Persistence of locality in systems with power-law interactions. *Phys. Rev. Lett.*, 113:030602, Jul 2014.

[88] Michael Foss-Feig, Zhe-Xuan Gong, Charles W. Clark, and Alexey V. Gorshkov. Nearly linear light cones in long-range interacting quantum systems. *Phys. Rev. Lett.*, 114:157201, Apr 2015.

[89] Andrew Y. Guo, Minh C. Tran, Andrew M. Childs, Alexey V. Gorshkov, and Zhe-Xuan Gong. Signaling and scrambling with strongly long-range interactions. *Phys. Rev. A*, 102:010401, Jul 2020.

[90] Minh C. Tran, Chi-Fang Chen, Adam Ehrenberg, Andrew Y. Guo, Abhinav Deshpande, Yifan Hong, Zhe-Xuan Gong, Alexey V. Gorshkov, and Andrew Lucas. Hierarchy of linear light cones with long-range interactions. *Phys. Rev. X*, 10:031009, Jul 2020.

[91] Zhe-Xuan Gong, Michael Foss-Feig, Fernando G. S. L. Brandão, and Alexey V. Gorshkov. Entanglement area laws for long-range interacting systems. *Phys. Rev. Lett.*, 119:050501, Jul 2017.

[92] Tomotaka Kuwahara and Keiji Saito. Area law of noncritical ground states in 1d long-range interacting systems. *Nat. Commun.*, 11(1):1–7, 2020.

[93] Jacob Biamonte. Lectures on quantum tensor networks. *arXiv preprint arXiv:1912.10049*, 2019.

[94] Jeffrey Pennington and Pratik Worah. The spectrum of the fisher information matrix of a single-hidden-layer neural network. *Adv. Neural Inf. Process. Syst.*, page 5410, 2018.

[95] Chae-Yeun Park and Michael J. Kastoryano. Geometry of learning neural quantum states. *Phys. Rev. Research*, 2:023232, May 2020.

[96] Mikhail Prokopenko, Joseph T. Lizier, Oliver Obst, and X. Rosalind Wang. Relating fisher information to order parameters. *Phys. Rev. E*, 84:041116, Oct 2011.

[97] Mohammad F. Maghrebi, Zhe-Xuan Gong, Michael Foss-Feig, and Alexey V. Gorshkov. Causality and quantum criticality in long-range lattice models. *Phys. Rev. B*, 93:125128, Mar 2016.

[98] Jorge Nocedal and Stephen Wright. *Numerical optimization*. Springer Science & Business Media, 2006.

[99] Scott Kirkpatrick. Frustration and ground-state degeneracy in spin glasses. *Phys. Rev. B*, 16:4630, Nov 1977.

[100] X. G. Wen and Q. Niu. Ground-state degeneracy of the fractional quantum hall states in the presence of a random potential and on high-genus riemann surfaces. *Phys. Rev. B*, 41:9377, May 1990.

[101] Xiao-Gang Wen. Topological orders in rigid states. *Int. J. Mod. Phys. B*, 4(02):239, 1990.

[102] Matthias Troyer and Uwe-Jens Wiese. Computational complexity and fundamental limitations to fermionic quantum monte carlo simulations. *Phys. Rev. Lett.*, 94:170201, May 2005.

[103] Michael Knap, Adrian Kantian, Thierry Giamarchi, Immanuel Bloch, Mikhail D. Lukin, and Eugene Demler. Probing real-space and time-resolved correlation functions with many-body ramsey interferometry. *Phys. Rev. Lett.*, 111:147205, Oct 2013.

[104] M Newman and G Barkema. *Monte carlo methods in statistical physics*. Oxford University Press: New York, USA, 1999.

[105] Davide Vodola, Luca Lepori, Elisa Ercolessi, Alexey V. Gorshkov, and Guido Pupillo. Kitaev chains with long-range pairing. *Phys. Rev. Lett.*, 113:156402, Oct 2014.

[106] For example, the canonical configuration of the bit array $(-1, 1, 1, -1, -1)$ should be $(-1, -1, -1, 1, 1)$ by our standard.

[107] Sebastian Fey, Sebastian C. Kapfer, and Kai Phillip Schmidt. Quantum criticality of two-dimensional quantum magnets with long-range interactions. *Phys. Rev. Lett.*, 122:017203, Jan 2019.

[108] Fangli Liu, Rex Lundgren, Paraj Titum, Guido Pagano, Jiehang Zhang, Christopher Monroe, and Alexey V. Gorshkov. Confined quasiparticle dynamics in long-range interacting quantum spin chains. *Phys. Rev. Lett.*, 122:150601, Apr 2019.

[109] Wen Lin Tan, Patrick Becker, F Liu, G Pagano, KS Collins, A De, L Feng, HB Kaplan, A Kyprianidis, R Lundgren, et al. Domain-wall confinement and dynamics in a quantum simulator. *Nat. Phys.*, 17(6):742–747, 2021.

[110] Chi-Fang Chen and Andrew Lucas. Finite speed of quantum scrambling with long range interactions. *Phys. Rev. Lett.*, 123:250605, Dec 2019.

[111] Xiao Chen and Tianci Zhou. Quantum chaos dynamics in long-range power law interaction systems. *Phys. Rev. B*, 100:064305, Aug 2019.

[112] David J. Luitz and Yevgeny Bar Lev. Emergent locality in systems with power-law interactions. *Phys. Rev. A*, 99:010105, Jan 2019.

[113] Nathan Goldman, Egidijus Anisimovas, Fabrice Gerbier, Patrik Öhberg, IB Spielman, and Gediminas Juzeliūnas. Measuring topology in a laser-coupled honeycomb lattice: from chern insulators to topological semi-metals. *New J. Phys.*, 15(1):013025, 2013.

[114] N. Read and Dmitry Green. Paired states of fermions in two dimensions with breaking of parity and time-reversal symmetries and the fractional quantum hall effect. *Phys. Rev. B*, 61:10267–10297, Apr 2000.

[115] G. Catalan, J. Seidel, R. Ramesh, and J. F. Scott. Domain wall nanoelectronics. *Rev. Mod. Phys.*, 84:119–156, Feb 2012.

[116] Takahiro Fukui, Yasuhiro Hatsugai, and Hiroshi Suzuki. Chern numbers in discretized brillouin zone: efficient method of computing (spin) hall conductances. *J. Phys. Soc. Japan*, 74(6):1674–1677, 2005.

[117] Chunlei Qu, Zhen Zheng, Ming Gong, Yong Xu, Li Mao, Xubo Zou, Guangcan Guo, and Chuanwei Zhang. Topological superfluids with finite-momentum pairing and majorana fermions. *Nat. Commun.*, 4(1):1–7, 2013.

[118] Bo Liu, Xiaopeng Li, Randall G. Hulet, and W. Vincent Liu. Detecting π-phase superfluids with p-wave symmetry in a quasi-one-dimensional optical lattice. *Phys. Rev. A*, 94:031602, Sep 2016.

[119] Lin Dong, Lei Jiang, and Han Pu. Fulde–ferrell pairing instability in spin–orbit coupled fermi gas. *New J. Phys.*, 15(7):075014, 2013.

[120] Yong Xu, Chunlei Qu, Ming Gong, and Chuanwei Zhang. Competing superfluid orders in spin-orbit-coupled fermionic cold-atom optical lattices. *Phys. Rev. A*, 89:013607, Jan 2014.

[121] Lianyi He, Hui Hu, and Xia-Ji Liu. Realizing fulde-ferrell superfluids via a dark-state control of feshbach resonances. *Phys. Rev. Lett.*, 120:045302, Jan 2018.

[122] Zhen-Fei Zheng, Guang-Can Guo, Zhen Zheng, and Xu-Bo Zou. Fulde–ferrell superfluids in spinless ultracold fermi gases. *New J. Phys.*, 20(6):063001, 2018.

[123] Edward Charles Titchmarsh, David Rodney Heath-Brown, Edward Charles Titchmarsh Titchmarsh, et al. *The theory of the Riemann zeta-function*. Oxford university press, 1986.

[124] Randall J LeVeque. *Finite difference methods for ordinary and partial differential equations: steady-state and time-dependent problems*. SIAM, 2007.

[125] Ruizhi Pan, Xuguang Yue, Xia Xu, Haichang Lu, and Xiaoji Zhou. Multiple photon-echo rephasing of coherent matter waves. *Phys. Lett. A*, 379(7):691–695, 2015.

[126] Hai-Chang Lu, Yue-Yang Zhai, Rui-Zhi Pan, and Shi-Feng Yang. An effective method of accelerating bose gases using magnetic coils. *Chin. Phys. B*, 23(9):093701, 2014.